

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: August 27, 2004, 16:27:32 ; Search time 474 Seconds
(without alignments)
2948.644 Million cell updates/sec

Title: US-10-081-775-2
Perfect score: 1718
Sequence: 1 MSTLGHNMESPHHTDVPDPS.....RKRVVRVFSQGMGKASE 329

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2.1/USPTO.spool/US10081775/runat 20082004.170212.11117/app_query.fasta_1.519
-DB=N Geneseq 29Jan04 -QFMT=fastap -SUFFIX=ring -MINMATCH=0.1 -LOOPCL=0
-LOOPTXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pct -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10081775.@CGN 1 1 470 @runat 20082004.170212.11117 -NCPU=6 -ICPU=3
-NO MMAP -LARGQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N Geneseq 29Jan04.*

- 1: Geneseq1980s.*
- 2: Geneseq1990s.*
- 3: Geneseq2000s.*
- 4: Geneseq2001as.*
- 5: Geneseq2001bs.*
- 6: Geneseq2002s.*
- 7: Geneseq2003as.*
- 8: Geneseq2003bs.*
- 9: Geneseq2003cs.*
- 10: Geneseq2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1718	100.0	1567	8 ACD91434	Acd91434 Human G-p
2	1691	98.4	990	5 AAS42339	Aas42339 Human cDN
3	1691	98.4	990	6 ABZ43046	Abz43046 Human GPC
4	1691	98.4	990	6 ABK37625	Abk37625 DNA encod
5	1691	98.4	998	6 ABQ88371	Abq88371 Human G p
6	1691	98.4	998	6 ABQ88372	Abq88372 Human G p
7	1691	98.4	1390	4 ADC86344	Adc86344 Human GPC
8	1649	96.0	963	4 AAH31850	Aah31850 Human oif

ALIGNMENTS

RESULT 1
ACD91434
ID ACD91434 standard; cDNA; 1567 BP.
XX
AC ACD91434;
XX
DT 22-SEP-2003 (first entry)
XX
DE Human G-protein coupled receptor HGPRBMV25 cDNA.
XX
KW Human; G-protein coupled receptor; HGPRBMV25; immune disease;
KW inflammatory disease; arthritis; asthma; AIDS; psoriasis;
KW graft-versus-host disease; systemic lupus erythematosus;
KW reproductive disorder; varicocele; orchitis; neural disorder;
KW Alzheimer's disease; Parkinson's disease; depression; schizophrenia;
KW cardiovascular disorder; hypertension; acute heart failure;
KW pulmonary disorder; endocrine disorder; obesity; diabetes; anorexia;
KW bone disorder; osteoporosis; pain; cancer; chromosome identification;
KW gene therapy; gene; ss.
XX
OS Homo sapiens.
XX
PN US2003060409-A1.
XX
PD 27-MAR-2003.
XX
PF 21-FEB-2002; 2002US-00081775.
XX
PR 21-FEB-2001; 2001US-0270134P.
XX
PR 27-MAR-2001; 2001US-0278952P.
XX
PA (RAMA/) RAMANATHAN C S.

Abk16633 Human G-c
Abk68612 Human DNA
Abz77974 Human G p
Adc86380 Human GPC
Abz77971 Human G p
Aah31721 Human oif
Aah32052 Human oif
Abz43027 Human GPC
Abk68587 Human DNA
Abk58762 Human G-p
Aaf88624 Human GCR
Adc86366 Human GCR
Aad12945 Human G-p
Abz52320 Human G-p
Abz81539 Human GPC
Aah31738 Human oif
Aas42265 Human cDN
Abz42976 Human GPC
Abk68617 Human DNA
Abk37551 DNA encod
Aba81537 Human GPC
Aas15906 DNA encod
Adc86330 Human GPC
Aba81538 Human GPC
Abq88366 Human G p
Adc86388 Human GPC
Aah32330 Human oif
Aas42259 Human cDN
Abz42968 Human GPC
Abk68605 Human DNA
Abk37545 DNA encod
Abz88354 Human G p
Abq88355 Human G p
Aad19141 Human G-p
Abz51301 Human REM
Aah31827 Human oif
Aas42338 Human cDN

Zozulya S.;

XX WPI; 2001-570867/64.
 DR P-PSDB; AAU24646.
 XX

Nucleic acids encoding human olfactory G protein-coupled receptors, useful for screening for compounds involved in olfactory sensation, where the compounds can be used in the food, pharmaceutical and cosmetic industries to customize odors.

XX Claim 1; Page 142; 319pp; English.
 XX

The invention relates to nucleic acids encoding human olfactory receptors, OR, (a G protein-coupled receptor, GPCR). The OR's specifically recognise molecules, odourants, that elicit specific olfactory sensation. The human olfactory receptors and polynucleotides encoding them are useful for screening a library of chemical compounds for compounds that are involved in olfactory sensation. Modulators of their activity are useful for pharmacological and genetic modulation of olfactory signalling pathways. Therefore, they can be used in the food, pharmaceutical and cosmetic industries to customise odours and fragrances. The present sequence encodes a human olfactory receptor of the invention

XX Sequence 990 BP; 188 A; 285 C; 221 G; 296 T; 0 U; 0 Other;
 SQ

Alignment Scores:
 Pred. No.: 4.97e-169 Length: 990
 Score: 1691.00 Matches: 325
 Percent Similarity: 99.09% Conservatives: 1
 Best Local Similarity: 98.78% Mismatches: 3
 Query Match: 98.43% Indels: 0
 DB: 5 Gaps: 0

US-10-081-775-2 (1-329) x AAS42339 (1-990)

QY 1 MetSerThrLeuGlyHisAsnMetGluSerProHisHisThrAspValAspProSer 20
 DB 1 ATGTCCAGCACTCTTGGCCCAACATGGAATCTCTTAATCACTCATGATGTTGACCCCTCT 60
 QY 21 ValPhePheLeuLeuGlyVileProGlyLeuGluGlnPheHisLeuTrpLeuSerLeuPro 40
 DB 61 GTCTTCTTCCTCCCTGGGATCCCAAGCTCGAACAATTTCAATTTGTGGCTCTCATCCCT 120
 QY 41 ValCysGlyLeuGlyThrAlaThrIleValGlyAsnIleThrIleLeuValValAla 60
 DB 121 GTGTGTGGCTTAGGCACAGCCACAAATGTGGCAATATACTATCTGTTGTTGTTGCC 180
 QY 61 ThrGluProValLeuHisLysProValTyrLeuPheLeuCysMetLeuSerThrIleAsp 80
 DB 181 ACTGAACAGTCTTGCAAGCCCTGTGTACCTTTTCTGTGCATGCTCTCAACCATCGAC 240
 QY 81 LeuAlaAlaSerValSerThrValProLysLeuLeuAlaIlePheTrpCysGlyAlaGly 100
 DB 241 TTGGCTGCCTCTGTCTCCACAGTTCACAGTCTCCCAAGCTACTGGCTATCTTCTGGTGAGCCGA 300
 QY 101 HisIleSerIleSerAlaCysLeuAlaHisMetPheIleHisAlaPheCysMetMet 120
 DB 301 CATATATCTGCCCTCTGCTGCTGGCCACAGATGTTCTTCAATTCATGCTTCTGCATGATG 360
 QY 121 GluSerThrValLeuLeuAlaMetAlaPheAspArgTyrValAlaIleCysHisProLeu 140
 DB 361 GATGCCACTGTGTACTGGCCATGGCTTTGTATCGCTACGTGGCCATCTGCCACCCACTC 420
 QY 141 ArgTyrAlaThrIleLeuThrAspThrIleIleAlaHisIleGlyValAlaValVal 160
 DB 421 CGCTATGCCACATCTCTCACTGACACCATCATTTGCCACATAGGGGTGGCAGTGTAGTG 480
 QY 161 ArgGlySerLeuLeuMetLeuProCysProPheIleGlyArgLeuAsnPheCysGln 180
 DB 481 CGAGGCTCCCTGCTCATGCTCCCATGCTCCCTTCTTATTTGGGGCTTTGAACTTCTGCCAA 540
 QY 181 SerHisValIleLeuHisThrTyrCysGluHisMetAlaValValLysLeuAlaCysGly 200

541 AGCATGTGTATCTTACACACGTTACTGTGAGCACATGCTGTGGTGAAGCTGGCCTGTGGA 600
 201 AspThrArgProAsnArgValTyrGlyLeuThrAlaAlaLeuLeuValIleGlyValAsp 220
 601 GACACAGGCTAACCGTGTGTATGGCTGACAGCTGCTGTGTGCTCATTTGGGTTGAC 660
 221 LeuPheCysIleGlyLeuSerTyrAlaLeuIleAlaGlnAlaValLeuArgLeuSerSer 240
 661 TTGTTTGGCATTTGCTCTCTCTATGCCCTAAGTGACACAGCTGCTCTTCCCTCTCATCC 720
 241 HisGluAlaArgSerLysAlaLeuGlyThrCysGlySerHisValCysValIleLeuIle 260
 721 CATGAAGCTCGGTCCMAAGCCCTTAGGACCTGTGGTTCCCATGCTGTGTGTCATCTCATC 780
 261 SerTyrThrProAlaLeuPheSerPhePheThrHisArgPheGlyHisHisValProVal 280
 781 TCTTATACACAGCCCTCTCTCTCTTTTACACACCGCTTTGGCCATCACGTTCCAGTC 840
 281 HisIleHisIleLeuLeuAlaAsnValTyrLeuLeuProProAlaLeuAsnProVal 300
 841 CATATTCACTTCTTTGGCCAATGTTTATCTGCTTTTGGCCACCTGCTTTAATCTCTGTG 900
 301 ValTyrGlyValLysThrLysGlnIleArgLysArgValValArgValPheGlnSerGly 320
 901 GTATATGGAGTTTAAGCAACAGATCCGTAAAGAGTTGTCAAGGTGTTTTCAAAGTGGG 960
 321 GlnGlyMetGlyVileLysAlaSerGlu 329
 961 CAGGGAATGGGCATCAAGGCATCTGAG 987

RESULT 3
 ABZ43046
 ID ABZ43046 standard; DNA; 990 BP.
 XX
 AC ABZ43046;
 XX
 DT 06-MAR-2003 (first entry)
 XX
 DE Human GPCR polynucleotide SEQ ID NO 353.
 XX
 KW Human; GPCR; G protein coupled receptor; signal transduction; olfactory;
 KW drug development; gustatory; taste; fragrance; gene; ds.
 XX Homo sapiens.
 XX WO200216548-A2.
 XX
 PD 28-FEB-2002.
 XX
 PF 30-JUL-2001; 2001WO-IB001446.
 XX
 PR 04-AUG-2000; 2000JP-00237818.
 PR 13-FEB-2001; 2001JP-00034434.
 XX
 PA (NISC-) JAPAN SCI & TECHNOLOGY CORP.
 XX
 PI Haga T, Takeda S, Mitaku S;
 XX
 DR WPI; 2002-304118/34.
 DR P-PSDB; ABP95772.
 XX
 PT Database global search for G protein-coupled receptors, proteins and
 PT encoded genes for studying in vivo signal transduction mechanism and
 XX identifying targets for drug development.
 PS Claim 9; SEQ ID NO 353; 97pp + Sequence Listing; Japanese.
 XX
 CC The invention relates to a method for screening G protein-coupled
 CC receptor (GPCR) genes (ABZ42870-ABZ43216) and/or GPCR proteins (ABP95596-
 CC ABP95942) by extracting open-reading frames containing 6-8 transmembrane
 CC domains with 250-1000 amino acid residues to give a gene homologous with
 CC a known GPCR gene. The receptor proteins and encoded genes are useful for
 CC studying in vivo signal transduction mechanism and identifying targets

Query Match:	98.43%	Indels:	0
DB:	6	Gaps:	0
US-10-081-775-2 (1-329) x ABK37625 (1-990)			
QY	1	MetSerSerThrLeuGlyHisAsnMetGluSerProHisHisThrAspValAspProSer	20
DB	1	ATGTCAGACACTCTTGGCCACACATGGAATCTCTTAATCACACTGATGTTGACCCCTTCT	60
QY	21	ValPhePheLeuGlyLeuProGlyLeuGluGlnPheHisLeuTrpLeuSerLeuPro	40
DB	61	GTCTTCTCTCTCTGGGCTCCAGGCTCTGGACAAATTTCAATTTGGGCTCTCACTCCCT	120
QY	41	ValCysGlyLeuGlyThrAlaThrLeuValGlyAsnLeuThrLeuValValAla	60
DB	121	GTGTGTGGCTTAGGCACACCAATTTGGGCAATATAACTATTCTGGTGTGTTGTC	180
QY	61	ThrGluProValLeuHisLysProValTyrLeuPheLeuCysMetLeuSerThrIleAsp	80
DB	181	ACTGAACCAAGCTTGGCAAGGCTGTGTACCTTTTCTGTGATGCTCTCAACCATCGAC	240
QY	81	LeuAlaAlaSerValSerThrValProLysLeuLeuAlaIlePheTrpCysGlyAlaGly	100
DB	241	TTGGCTGCTCTGTCTCCACAGTCCCAAGCTACTGGCTATCTTCTCATTCATGCTTCTG	300
QY	101	HisIleSerAlaSerAlaCysLeuAlaHisMetPhePheIleHisAlaPheCysMetMet	120
DB	301	CATATATCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	360
QY	121	GluSerThrValLeuLeuAlaMetAlaPheAspArgTyrValAlaIleCysHisProLeu	140
DB	361	GAGTCCACTGTGCTACTGGCCATGGCTTTGATGCTAGTGGCCATCTGCCACCCATC	420
QY	141	ArgTyrAlaThrIleLeuThrAspThrIleLeuAlaHisIleGlyValAlaValVal	160
DB	421	CGCTATGCCAATCTCTCACTGACACCATCAATGCCCATAGAGGCTGGCAGCTGTAGT	480
QY	161	ArgGlySerLeuLeuMetLeuProCysProPhePheIleGlyArgLeuAsnPheCysGln	180
DB	481	CGAGGCTCCCTGCTCATGCTCCCATGTCCTTCTTATGTCGCTGTTGAACTTCTGCCA	540
QY	181	SerHisValIleLeuHisThrTyrCysGluHisMetAlaValValLysLeuAlaCysGly	200
DB	541	AGCCATGTGATCTCTACACACACTGCTGAGCAGCATGCTGCTGAGCTGGCTGTGGA	600
QY	201	AspThrArgProAsnArgValTyrGlyLeuThrAlaAlaLeuLeuValIleGlyValAsp	220
DB	601	GACACCAAGGCTTAACCGTGTGTATGGGCTGACAGCTGCTGCTGCTGCTGCTGCTG	660
QY	221	LeuPheCysIleGlyLeuSerTyrAlaLeuIleAlaGlnAlaValLeuArgLeuSerSer	240
DB	661	TTGTTTGTGATGGTCTCTCTATGCCCTAAGTGCAAGCTGCTGCTGCTGCTGCTGCT	720
QY	241	HisGluAlaArgSerLysAlaLeuGlyThrCysGlySerHisValCysValIleLeuIle	260
DB	721	CATGAAGCTCGTCCCAAGGCTTAGGACCTGTGTTTCCCATGCTGCTGCTGCTGCTG	780
QY	261	SerTyrThrProAlaLeuPheSerPhePheThrHisArgPheGlyHisHisValProVal	280
DB	781	TCCTTATACACAGCCCT	840
QY	281	HisIleHisIleLeuLeuAlaAsnValTyrLeuLeuLeuProProAlaLeuAsnProVal	300
DB	841	CATATTACATCTTTTGGCCATGTTTATCTGCTTTTGGCCATGCTGCTGCTGCTGCTG	900
QY	301	ValTyrGlyValLysThrLysGlnIleArgLysArgValValArgValPheGlnSerGly	320
DB	901	GTATATGGAGTTAAGACCAACAGATCCGTAAGAGTTGTGAGGTTGTTCAAAGTGGG	960
QY	321	GlnGlyMetGlyIleLysAlaSerGlu	329
DB	961	CAGGGAATGGCATCAAGGCATCTGAG	987

RESULT 5
ABQ88371
ID ABQ88371 standard; cDNA; 998 BP.
XX
AC ABQ88371;
XX
DT 23-SEP-2002 (first entry)
XX
Human G protein coupled receptor cDNA SEQ ID NO:35.
XX
Human; G protein coupled receptor; GPCR; GPCRX; neuroprotective;
XX
nootropic; anti-HIV; antiasthmatic; antiarteriosclerotic; cyostatic;
XX
immunomodulator; antiinflammatory; antidiabetic; anorectic; haemostatic;
XX
antibacterial; fungicide; protozoal; virucide; nephrotropic; osteopathic;
XX
cardiac; antulcer; antiallergic; hepatotropic; antiparkinsonian; HIV;
XX
vaccine; gene therapy; cell signal processing; cardiomyopathy; diabetes;
XX
metabolic pathway modulation; atherosclerosis; cancer; obesity; asthma;
XX
infection; Parkinson's disease; osteoporosis; Crohn's disease; ulcer;
XX
allergy; cirrhosis; glomerulonephritis; stroke; haematopoietic disorder;
XX
systemic lupus erythematosus; gene; ss.
XX
Homo sapiens.
XX
WO200250276-A2.
XX
27-JUN-2002.
XX
18-DEC-2001; 2001WO-US049347.
XX
18-DEC-2000; 2000US-0256635P.
XX
21-DEC-2000; 2000US-0257876P.
XX
04-JAN-2001; 2001US-0259743P.
XX
10-JAN-2001; 2001US-0260718P.
XX
12-JAN-2001; 2001US-0261498P.
XX
24-JAN-2001; 2001US-0263689P.
XX
08-FEB-2001; 2001US-0267454P.
XX
22-FEB-2001; 2001US-0271021P.
XX
14-MAR-2001; 2001US-0275946P.
XX
23-MAR-2001; 2001US-0278150P.
XX
18-APR-2001; 2001US-0284591P.
XX
23-APR-2001; 2001US-0285718P.
XX
19-JUN-2001; 2001US-0299327P.
XX
16-AUG-2001; 2001US-0312902P.
XX
(CURA-) CURAGEN CORP.
XX
Li L, Padigar M, Ballinger RA, Kekuda R, Colman SD, Sciore P;
XX
Smithson G, Peyman JA, Macdougall JR, Stone D, Vernet CAM, Shenoy S;
XX
Gunther E, Millet I, Tchernev VT, Anderson D, Gusev V, Malyankar UM;
XX
Zhong H, Ellerman KE, Wolenc A;
XX
WPI; 2002-557660/59.
XX
P-PSDB; ABP51577.
XX
New isolated human G-protein coupled receptor X (GPCRX) polypeptide,
XX
useful for treating or preventing GPCR-associated disorders e.g.
XX
diabetes, atherosclerosis, cancer or obesity.
XX
Claim 8; Page 88; 354pp; English.
XX
ABQ88354 to ABQ88417 represent human G protein coupled receptor (GPCR)
XX
cDNA sequences, and ABP51560 to ABP51624 represent human GPCR proteins
XX
from the present invention. GPCR sequences can have neuroprotective,
XX
nootropic, anti-HIV, antiasthmatic, antiarteriosclerotic, cyostatic,
XX
immunomodulator, antiinflammatory, antidiabetic, anorectic, haemostatic,
XX
antibacterial, fungicide, protozoal, virucide, nephrotropic, osteopathic,
XX
cardiac, antulcer, antiallergic, hepatotropic and antiparkinsonian
XX
activities, and can be used in vaccines and gene therapy. GPCR proteins,
XX
nucleic acid molecules, and antibodies from the present invention can be
XX
used for manufacturing a medicament for treating or preventing a GPCR-
XX
associated disorder or syndrome related to cell signal processing and
XX
metabolic pathway modulation, such as cardiomyopathy, atherosclerosis,
XX
diabetes, cancer, obesity, infections (bacterial, fungal, protozoal or

CC viral), HIV, asthma, Parkinson's disease, osteoporosis, Crohn's disease,
CC ulcers, allergies, cirrhosis, glomerulonephritis, stroke, systemic lupus
CC erythematous, or haematopoietic disorders. Anti-GPCR antibodies can be
CC used diagnostically to monitor protein levels in tissues as part of a
CC clinical testing procedure such as in determining the efficacy of a given
CC treatment regimen. ABQ88418 to ABQ88639 represent PCR primers and probes
CC for the human GPCRs of the present invention
XX
SQ Sequence 998 BP; 190 A; 287 C; 223 G; 298 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 5.03e-169 Length: 998
Score: 1691.00 Matches: 325
Percent Similarity: 99.09% Conservative: 1
Best Local Similarity: 98.78% Mismatches: 3
Query Match: 98.43% Indels: 0
DB: 6 Gaps: 0

US-10-081-775-2 (1-329) x ABQ88371 (1-998)

QY 1 MetSerThrLeuGlyHisAsnMetGluSerProHisHisThrAspValAspProSer 20
DB 2 ATGTCCAGCCTCTTGCCACACATGCAATCTCTAATCACTGATGTTGACCTTCT 61
QY 21 ValPhePheLeuGlyIleProGlyLeuGluGlnPheHisLeuTrpLeuSerLeuPro 40
DB 62 GTCTTCTCTCTCTGGGATCCCAAGTCTGGAAACAATTTCAATTGTGGCTCTCACTCCCT 121
QY 41 ValCysGlyLeuGlyThrAlaThrIleValGlyAsnIleThrIleLeuValValAla 60
DB 122 GTGTGTGGCTTAGGCACAGCCACAAATGTGGCATATACTATTCTGTGTTGTTGCC 181
QY 61 ThrGluProValLeuHisLysProValTyrLeuPheLeuCysMetLeuSerThrIleAsp 80
DB 182 ACTGAACAGCTTGTGCACAGCTGTGTACCTTTTCTGTGATCTCTCAACCATCCAC 241
QY 81 LeuAlaAlaSerValSerThrValProLysLeuLeuAlaIlePheTrpCysGlyValAlaGly 100
DB 242 TTGGCTGCTCTGTCTCCACAGTCTCCCAAGTACTGGGTATCTTCTGTGTGGAGCCGGA 301
QY 101 HistLeSerAlaSerAlaCysLeuAlaHisMetPhePheIleHisAlaPheCysMetMet 120
DB 302 CATATATCTGCTCTGCTGCTGCGACAGATGTTCTTCAATCATGCTCTCTGCATGATG 361
QY 121 GluSerThrValLeuLeuAlaMetAlaPheAspArgTyrValAlaIleCysHisProLeu 140
DB 362 GAGTCCACTGTGCTACTGGCCATGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 421
QY 141 ArgTyrAlaThrIleLeuThrAspThrIleAlaHisIleGlyValAlaAlaValVal 160
DB 422 CGCTATGCCACAACTCTCACTGACACCATCATTTGCCCATAGGGGTGGCAGCTGTAGTG 481
QY 161 ArgGlySerLeuLeuMetLeuProCysProPhePheIleGlyArgLeuAsnPheCysGln 180
DB 482 CAGAGCTCCTGCTCATGCTCCCATGCTCCCTTCTTATTTGGGGTTTGAATCTTGCCAA 541
QY 181 SerHisValIleLeuHisThrTyrCysGluHisMetAlaValValLysLeuAlaCysGly 200
DB 542 AGCCATGTATCTCATACACAGTACTGTGAGCAGCATGGCTGTGTGAGCTGGCCCTGTGA 601
QY 201 AspThrArgProAsnArgValTyrGlyLeuThrAlaAlaLeuLeuValIleGlyValAsp 220
DB 602 GACACAGGCCCTAACCGTGTATGGCGTGCACAGCTGCTGCTGCTGCTGCTGCTGCTGCT 661
QY 221 LeuPheCysIleGlyLeuSerTyrAlaLeuIleAlaGlnAlaValLeuArgLeuSerSer 240
DB 662 TTGTTTGTGATTTGTTCTCTCTCTATGCTTAAAGTGCACAGCTGCTGCTGCTGCTGCT 721
QY 241 HisGluAlaArgSerLysAlaLeuGlyThrCysGlySerHisValCysValIleLeuIle 260
DB 722 CATGAAGCTCGGTCCCAAGGCCCTTAGGACCTGTGTGTTCCCATGCTGTGTGTCATCCTCATC 781
QY 261 SerTyrThrProAlaLeuPheSerPheThrHisArgPheGlyHisValProVal 280

DB 782 TCCTATACACAGCCCT 841
QY 281 HisIleHisIleLeuLeuAlaAsnValTyrLeuLeuLeuProAlaLeuAsnProVal 300
DB 842 CATATTCACATCTCTTTGGCCAATGTTATCTGCTTTGGCAGCTGCTCTTAATCTCTGTG 901
QY 301 ValTyrGlyValLysThrLysGlnIleArgLysArgValValArgValPheGlnSerGly 320
DB 902 GTATATGGAGTTAAGACCAACACAGATCCGTAAGAAGTGTCTCAGGGGTGTTTCAAGTGGG 961
QY 321 GlnGlyMetGlyIleLysAlaSerGlu 329
DB 962 CAGGGAATGGGCATCAAGGCATCTGAG 988
AC ABQ88372 standard; cDNA; 998 BP.
XX ABQ88372;
XX AC
XX 23-SEP-2002 (first entry)
XX DE
XX Human G protein coupled receptor cDNA SEQ ID NO:37.
XX Human; G protein coupled receptor; GPCR; GPCR; neuroprotective;
KW norepinephrine; anti-HIV; antiasthmatic; antiarteriosclerotic; cytostatic;
KW immunomodulator; antiinflammatory; antidiabetic; anorectic; haemostatic;
KW antibacterial; fungicide; protozoal; virucide; hepatotropic; antiparkinsonian; HIV;
KW cardiant; antitumor; antiallergic; hepatotropic; cardiomyopathy; diabetes;
KW vaccine; gene therapy; cell signal processing; cancer; obesity; asthma;
KW metabolic pathway modulation; atherosclerosis; Crohn's disease; ulcer;
KW infection; Parkinson's disease; osteoporosis; stroke; haematopoietic disorder;
KW allergy; cirrhosis; glomerulonephritis; gene; ss.
XX Homo sapiens.
XX OS
XX WO200250276-A2.
XX PD
XX 27-JUN-2002.
XX PF
XX 18-DEC-2001; 2001WO-US049347.
XX PR
XX 18-DEC-2000; 2000US-0256663P.
XX PR
XX 21-DEC-2000; 2000US-0257876P.
XX PR
XX 04-JAN-2001; 2001US-0259743P.
XX PR
XX 10-JAN-2001; 2001US-0260718P.
XX PR
XX 12-JAN-2001; 2001US-0261498P.
XX PR
XX 24-JAN-2001; 2001US-0263689P.
XX PR
XX 08-FEB-2001; 2001US-0267464P.
XX PR
XX 22-FEB-2001; 2001US-0271021P.
XX PR
XX 14-MAR-2001; 2001US-0275946P.
XX PR
XX 23-MAR-2001; 2001US-0278150P.
XX PR
XX 18-APR-2001; 2001US-0284591P.
XX PR
XX 23-APR-2001; 2001US-0285718P.
XX PR
XX 19-JUN-2001; 2001US-0299327P.
XX PR
XX 16-AUG-2001; 2001US-0312902P.
XX (CURA-) CURAGEN CORP.
XX PA
XX Li L, Padigar M, Ballinger RA, Kekuda R, Colman SD, Sciore P;
PI Smithson G, Peyman JA, Macdougall JR, Stone D, Vernet CAM, Shenoy S;
PI Gunther E, Millet I, Tchernev VT, Anderson D, Gusev V, Malyankar UM;
PI Zhong H, Ellerman KE, Wolenc A;
XX WPI; 2002-557660/59.
DR P-PSDB; ABP51578.
XX
XX New isolated human G-protein coupled receptor X (GPCRX) polypeptide,
XX useful for treating or preventing GPCR-associated disorders e.g.
XX diabetes, atherosclerosis, cancer or obesity.

PS Claim 8; Page 88; 354pp; English.

CC ABQ88354 to ABQ88417 represent human G protein coupled receptor (GPCR) CDNA sequences, and ABP51560 to ABP51624 represent human GPCR proteins from the present invention. GPCR sequences can have neuroprotective, CC nootropic, anti-HIV, antiasthmatic, antiarteriosclerotic, cytostatic, CC immunomodulator, antiinflammatory, antidiabetic, anorectic, haemostatic, CC antibacterial, fungicide, protozoal, virucide, nephrotropic, osteopathic, CC cardiant, antiulcer, antiallergic, hepatotropic and antiparkinsonian CC activities, and can be used in vaccines and gene therapy. GPCR proteins, CC nucleic acid molecules, and antibodies from the present invention can be CC used for manufacturing a medicament for treating or preventing a GPCR- CC associated disorder or syndrome related to cell signal processing and CC metabolic pathway modulation, such as cardiomyopathy, atherosclerosis, CC diabetes, cancer, obesity, infections (bacterial, fungal, protozoal or CC viral), HIV, asthma, Parkinson's disease, osteoporosis, Crohn's disease, CC ulcers, allergies, cirrhosis, glomerulonephritis, stroke, systemic lupus CC erythematosus, or haematopoietic disorders. Anti-GPCR antibodies can be CC used diagnostically to monitor protein levels in tissues as part of a CC clinical testing procedure such as in determining the efficacy of a given CC treatment regimen. ABQ88418 to ABQ88639 represent PCR primers and probes CC for the human GPCRs of the present invention

XX

XX Sequence 998 BP; 190 A; 287 C; 223 G; 298 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 5.03e-169 Length: 998
Score: 1691.00 Matches: 325
Percent Similarity: 99.09% Conservative: 1
Best Local Similarity: 98.78% Mismatches: 3
Query Match: 98.43% Indels: 0
DB: 6 Gaps: 0

US-10-081-775-2 (1-329) x ABQ88372 (1-998)

Qy 1 MetSerThrLeuGlyHisAsnMetGluSerProHisHisThrAspValAspProSer 20
Db 2 ATGTCCAGCACTCTGTGGCCACCAACATGGAATCTCTAATACACATGATGTTGACCCCTCT 61
Qy 21 ValPhePheLeuLeuGlyIleProGlyLeuGluGlnPheHisLeuTyrLeuSerLeuPro 40
Db 62 GTCTTCTCTCTGGGCATCCAGCTGTGGACAAATTCATTGTGGCTCTCACTCCCT 121
Qy 41 ValCysGlyLeuGlyThrAlaThrIleValGlyAsnIleThrIleLeuValValAla 60
Db 122 GTGTGTGGCTTAGGCACAGCCACAAATGTGGCAATATACTATCTCTGTGTGTGGCC 181
Qy 61 ThrGluProValLeuHisLysProValTyrLeuPheLeuCysMetLeuSerThrIleAsp 80
Db 182 ACTGAACCACTGTGCAAGCCCTGTGTACCTTTTCTGTGCATGCTCTCAACCATCGAC 241
Qy 81 LeuAlaAlaSerValSerThrValProLysLeuLeuAlaIlePheTyrCysGlyAlaGly 100
Db 242 TTGGCTGCCTCTGTCTCCACAGTTCACAGCTACTGGCTATCTTCTGTGGTGTGGAGCCGA 301
Qy 101 HisIleSerAlaSerAlaCysLeuAlaHisMetPheIleHisAlaPheCysMetMet 120
Db 302 CATATATCTGCCTCTGCCTGGCCGACACATGTTCTTCATTATGCTTCTGCATGATG 361
Qy 121 GluSerThrValLeuLeuAlaMetAlaPheAspArgTyrValAlaIleCysHisProLeu 140
Db 362 GAGTCCACTGTGTACTGGCCATGGCTTGTGATCGCTAGTGGCCATCTGCCACCCATC 421
Qy 141 ArgTyrAlaThrIleLeuThrAspThrIleIleAlaHisIleGlyValAlaValVal 160
Db 422 CGTATATGCCACAACTCTCACTGACACATCATTTGCCCATATAGGGGTGGCAGTGTAGTG 481
Qy 161 ArgGlySerLeuLeuMetLeuProCysProPhePheIleGlyArgLeuAsnPheCysGln 180
Db 482 CGAGGCTCCCTGCTATGCTCCATGCTCCATGCTCTTCTTATTTGGCGTTTGAACTTCTGCCAA 541
Qy 181 SerHisValIleLeuHisThrTyrCysGluHisMetAlaValValLysLeuAlaCysGly 200

Db 542 AGCCATGTGATCTCTACACAGTACTGTGAGCACATGCTGTGCTGAAGCTGGCTGTGGA 601
Qy 201 AspThrArgProAsnArgValTyrGlyLeuThrAlaAlaLeuLeuValIleGlyValAsp 220
Db 602 GACACAGGCGCTAACCGTGTGTATGGCTGACAGCTGCACTGTGTGTCATTTGGGTTGAC 661
Qy 221 LeuPheCysIleGlyLeuSerTyrAlaLeuIleAlaGlnAlaValLeuArgLeuSerSer 240
Db 662 TTGTTTGTGATGGTCTCTCTATGCCCTAAGTGGACAGCTGCTTGGCCTCTCATCC 721
Qy 241 HisGluAlaArgSerLysAlaLeuGlyThrCysGlySerHisValCysValIleLeuIle 260
Db 722 CATGAAGCTCGGTGCAAGGCCCTAGGGACCTGTGGTTCCCATGCTCTGTGTCATCCTCATC 781
Qy 261 SerTyrThrProAlaLeuPheSerPhePheThrHisArgPheGlyHisHisValProVal 280
Db 782 TCTTATACACAGCCCTCTTCTCTCTTTTACACACCGCTTTGGCCATCACGTTCCAGTC 841
Qy 281 HisIleHisIleLeuLeuAlaAsnValTyrLeuLeuLeuProAlaLeuAsnProVal 300
Db 842 CATATTCAATTCTTTTGGCCAAATGTTTATCTGTCTTTGCCACCTGCTTTAATCTGTG 901
Qy 301 ValTyrGlyValLysThrLysGlnIleArgLysArgValValArgValPheGlnSerGly 320
Db 902 GTATATGGAGTTAAGACCAACAGATCCGTAAACAGATTGTCAAGGCTGTTTCAAAAGTGGG 961
Qy 321 GlnGlyMetGlyIleLysAlaSerGlu 329
Db 962 CAGGGAATGGGCATCAAGGCATCTGAG 988
RESULT 7
ADC86344
ID ADC86344 standard; DNA; 1390 BP.
XX
AC ADC86344;
XX
DT 01-JAN-2004 (first entry)
XX
DE Human GPCR gene SEQ ID NO:797.
XX
KW ds; gene; human; GPCR;
XX
OS guanosine triphosphate-binding protein coupled receptor; gene therapy.
XX
FN Homo sapiens.
XX
EP EPI270724-A2.
XX
PD 02-JAN-2003.
XX
PF 18-JUN-2002; 2002EP-00013517.
XX
PR 18-JUN-2001; 2001JP-00246789.
XX
PA (NAAD-) NAT INST ADVANCED IND SCI & TECHNOLOGY.
XX
PA (ADSC-) CENT ADVANCED SCI & TECHNOLOGY INCUBATIO.
XX
PI Suwa M, Asai K, Akiyama Y, Aburatani H;
XX
WI WPI; 2003-315783/31.
XX
P-PSDB; ADC86345.
XX
PT New polynucleotide, useful for preparing a composition for treating a
XX patient in need of increased or suppressed activity or expression of the
XX guanosine triphosphate-binding protein coupled receptor.
XX
PS Claim 1; SEQ ID NO 797; 28pp; English.
XX
CC The invention relates to a novel polynucleotide encoding a guanosine
XX triphosphate-binding protein coupled receptor (GPCR). A polynucleotide of
XX the invention may have a use in gene therapy. The polynucleotide and
XX polypeptide are useful for preparing a composition for treating a patient
XX in need of increased or suppressed activity or expression of the
XX guanosine triphosphate-binding protein coupled receptor. The

CC polynucleotide sequences shown in ADC85548-ADC87616 encode GPCR's of the
 CC invention.
 XX
 SQ Sequence 1390 BP; 317 A; 353 C; 293 G; 427 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 8,03e-169 Length: 1390
 Score: 1691.00 Matches: 325
 Percent Similarity: 99.09% Conservativity: 1
 Best Local Similarity: 98.78% Mismatches: 3
 Query Match: 98.43% Indels: 0
 DB: 9 Gaps: 0
 US-10-081-775-2 (1-329) x ADC86344 (1-1390)
 QY 1 MetSerThrLeuGlyHisAsnMetGluSerProHisHisThrAspValAspProSer 20
 Db
 201 ATGTCCAGCACTCTTGGCCACACATGGATCTCTTAATCACACTGATGTTGACCTTCT 260
 QY 21 ValPheLeuLeuGlyIleProGlyLeuGluGlnPheHisLeuTrpLeuSerLeuPro 40
 Db
 261 GTCTTCTCTCTCGGCGCATCCAGGCTGGACAAATTTCAATTTGTGGCTCTCACTCCCT 320
 QY 41 ValCysGlyLeuGlyThrAlaThrIleValGlyAsnIleThrIleLeuValValAla 60
 Db
 321 GTGTGTGGCTTAGGCACACACACATTTGGGGCAATATACTATTTCTGGTTGTGGCC 380
 QY 61 ThrGluProValLeuHisLeuProValTrpLeuPheLeuCysMetLeuSerThrIleAsp 80
 Db
 381 ACTGACCAAGCTTTCACACAGCTGTGTACCTTTTCTGTGATGCTCTCAACATCGAC 440
 QY 81 LeuAlaAlaSerValSerThrValProLysLeuLeuAlaIlePheTrpCysGlyValaGly 100
 Db
 441 TTGGCTGCTCTGTCTCCACAGCTTCCCAAGCTACTGGCTATCTTCTGGTGTGGAGCCGGA 500
 QY 101 HisIleSerAlaSerAlaCysLeuAlaHisMetPhePheIleHisAlaPheCysMetMet 120
 Db
 501 CATATATCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 560
 QY 121 GluSerThrValLeuLeuAlaMetAlaPheAspArgTrpValAlaIleCysHisProLeu 140
 Db
 561 GAGTCCACTGTCTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 620
 QY 141 ArgTrpAlaThrIleLeuLeuThrAspThrIleIleAlaHisIleGlyValAlaValVal 160
 Db
 621 CGCTATGCAATCTCACTGACACCATCATTTGCCACATAGGGGTGGCAGCTGTAGTG 680
 QY 161 ArgGlySerLeuLeuMetLeuProCysProPhePheIleGlyValArgLeuAsnPheCysGln 180
 Db
 681 CGAGGCTCCCTGCTCATGCTCCATGCTCCATGCTCCATGCTCCATGCTCCATGCTCCATGCT 740
 QY 181 SerHisValIleLeuHisThrTrpCysGluHisMetAlaValValLysLeuAlaCysGly 200
 Db
 741 ACCCATGTGATCTTACACAGCTACTGTGAGCAGCATGGCTGTGGTGAAGCTGGGCTGTGA 800
 QY 201 AspThrArgProAsnArgValTrpGlyLeuThrAlaAlaLeuLeuValIleGlyValAsp 220
 Db
 801 GACACAGGCTTAACCGTGTGTATGGGCTGACAGCTGCACTGTGTGTGTGTGTGTGTGTGTG 860
 QY 221 LeuPheCysIleGlyLeuSerTrpAlaLeuIleAlaGlnAlaValLeuArgLeuSerSer 240
 Db
 861 TTGTTTTGTGATTTGGTCTCTCTGATGCTTAAAGTGCACAGCTGCTCTGCTCTCATCC 920
 QY 241 HisGluAlaArgSerLysAlaLeuGlyThrCysGlySerHisValCysValIleLeuIle 260
 Db
 921 CATGAGCTCGTCCAGGCTTAGGACCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCT 980
 QY 261 SerTrpThrProAlaLeuPheSerPhePheThrHisArgPheGlyHisHisValProVal 280
 Db
 981 TCTTATACACAGGCTCTTCT 1040
 QY 281 HisIleHisIleLeuLeuAlaAsnValTrpLeuLeuLeuProProAlaLeuAsnProVal 300

Db 1041 CATATTACATCTTTTGGCCAAATTTATCTCTTTGGCCACTCTCTTATCTCTGTG 1100
 QY 301 ValTrpGlyValLysThrLysGlnIleArgLysArgValValArgValPheGlnSerGly 320
 Db 1101 GTATATGGAGTTAAGACCAACAGATCCGTAAGAAGTTGTCAAGGTGTTTCAAGTGGG 1160
 QY 321 GlnGlyMetGlyIleLysAlaSerGlu 329
 Db 1161 CAGGGAATGGCATCAAGGCATCTGAG 1187
 RESULT 8
 AAH31850
 ID AAH31850 standard; DNA; 963 BP.
 XX AAH31850;
 AC AAH31850;
 XX 30-JUL-2001 (first entry)
 DT
 XX Human olfactory receptor polynucleotide, SEQ ID NO: 423.
 DE
 XX Human; olfactory receptor; OR; primary scent determination;
 KW secondary scent determination; polypeptide library; odour receptor;
 KW scent profile; scent fingerprint; scent representation; ds.
 XX Homo sapiens.
 OS
 XX WO200127158-A2.
 PN
 XX 19-APR-2001.
 PD
 XX 06-OCT-2000; 2000MO-US027582.
 PF
 XX 08-OCT-1999; 99US-0158615P.
 PR 24-FEB-2000; 2000US-0184809P.
 XX (DIGI-) DIGISCENTS.
 PA (YEDA) YEDA RES & DEV CO LTD.
 XX Bellensohn J, Smith D, Lancet D, Glusman G, Fuchs T, Yanai I;
 PI WPI; 2001-290713/30.
 XX New polynucleotides which encode polypeptides involved in olfactory
 PT sensation for identifying olfactory agonists and antagonists.
 PS Claim 8; Page 349; 1857pp; English.
 XX The present sequence is one of a number of isolated polynucleotides which
 CC encode polypeptides involved in olfactory sensation. The polynucleotides
 CC can be used in screening for olfactory agonists and antagonists. The
 CC methods allow for the determination of primary scents and the
 CC identification of the odour receptors used to detect these primary
 CC scents. The methods also enable determination of secondary scents and the
 CC identification of combinations of odour receptors that are involved in
 CC detecting such secondary scents. This enables the construction of a scent
 CC representation (also called a scent fingerprint or scent profile), which
 CC may be used to re-create and edit scents. Libraries of olfactory
 CC receptors are useful for determining the interaction pattern of a
 CC composition with the receptors, and can be used for determining
 CC differences in the olfactory faculties of different individuals
 XX
 SQ Sequence 963 BP; 181 A; 276 C; 216 G; 290 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 1.33e-164 Length: 963
 Score: 1649.00 Matches: 317
 Percent Similarity: 99.07% Conservativity: 1
 Best Local Similarity: 98.75% Mismatches: 3
 Query Match: 95.98% Indels: 0
 DB: 4 Gaps: 0
 US-10-081-775-2 (1-329) x AAH31850 (1-963)

QY	9	MecGlusSerProHisHisThrAspValAspProSerValPhePheLeuLeuGlyIlePro	20
DB	1	ATCGAAATCTCTAATACACACTGATGTGACCCCTTCGTCTCTCTCTCTGGGCATCCCA	60
QY	29	GlyLeuGluGlnPheHisLeuTrpLeuSerLeuProValCysGlyLeuGlyThrAlaThr	48
DB	61	GGTCTGGACAAATTCATTTGTGGCTCTCACTCCCTGTGTGTGGCTTAGGCACAGCCACA	120
QY	49	IleValGlyAsnIleThrIleLeuValValAlaThrGluProValLeuHisLysPro	68
DB	121	ATTGTGGGCAATATAACTATTCCTGGTGTGTGTGCCACTGAACCACTGTGCACAAAGCT	180
QY	69	ValTyrLeuPheLeuCysMetLeuSerThrIleAspLeuAlaSerValSerThrVal	88
DB	181	GGTAGACCTTTTCTGTGCACTGTCTCAACCACTCACTGGCTGGCTCTGTCTCCACAGTT	240
QY	89	ProLysLeuLeuAlaIlePheTrpCysGlyAlaGlyHisIleSerAlaSerAlaCysLeu	108
DB	241	CCCAAGCTACTGGCTATCTTTCTGGTGTGGAGCGGACATATATCTGCCTCTGCCTGCCTG	300
QY	109	AlaHisMetPhePheIleHisAlaPheCysMetMetGluSerThrValLeuLeuAlaMet	128
DB	301	GCACAGATGTTCTTCATTCATGCCCTCTGCATGTGAGTCCACTGTGCTACTGGCCATG	360
QY	129	AlaPheAspArgTyrValAlaIleCysHisProLeuArgTyrAlaThrIleLeuThrAsp	148
DB	361	GCCTTGTATCGGTACGTGGCCATCTGCCACCACCTCCGCTATGCCACAATCTCACTGAC	420
QY	149	ThrIleIleAlaHisIleGlyValAlaValAlaValArgGlySerLeuLeuMetLeuPro	168
DB	421	ACCATCATTTGCCACATAGGGGTGCACGTGTAGTGCAGGCTCCCTGCTCATGTGCCA	480
QY	169	CysProPhePheIleGlyArgLeuAsnPheCysGlnSerHisValIleLeuHisThrTyr	188
DB	481	TGTCCTCTTCCTATTATGGCGCTTTGAACCTCTGCCAAAGCCATGTGAATCTACACAGTAC	540
QY	189	CysGluHisMetAlaValValValValAlaCysGlyAspThrArgProAsnArgValTyr	208
DB	541	TGTGAGCACATGGCTGTGGTGAAGCTGGCCCTGTGGAGACACAGGCCCTAACCGTGTGTAT	600
QY	209	GlyLeuThrAlaAlaLeuLeuValIleGlyValAspLeuPheCysIleGlyLeuSerTyr	228
DB	601	GGGCTGACAGCTGACCTGTTGGTTCATTGGGGTTGACTTGTTCATGTTGGTCTCTCTAT	660
QY	229	AlaLeuIleAlaGlnAlaValLeuArgLeuSerSerHisGluAlaArgSerLysAlaLeu	248
DB	661	GCCCTAAGTGCACAAGCTGTCCTTCGGCTCTCATCCCATGAAGCTCGTCCAAAGGCCCTA	720
QY	249	GlyThrCysGlySerHisValCysValIleLeuIleSerTyrThrProAlaLeuPheSer	268
DB	721	GGGACCTGTGGTTCCTCATGTCTGTGTCATCTCATCTTATACACAGGCCCTCTTCTCC	780
QY	269	PhePheThrHisArgPheGlyHisHisValProValHisIleHisIleLeuLeuAlaAsn	288
DB	781	TTTTTTACACACCGCTTTGGCCATCAGTTCAGTCCATATTCACATTCCTTTTGGCAAT	840
QY	289	ValTyrLeuLeuLeuProProAlaLeuAsnProValValTyrGlyValLysThrLysGln	308
DB	841	GTTTATCTGCTTTTGGCCACTGCTCTTAATCCCTGTGTGTATATGGAGTTAAGACCAACAG	900
QY	309	IleArgLysArgValValArgValPheGlnSerGlyGlnGlyMetGlyIleLysAlaSer	328
DB	901	ATCCGTAAAGAGTGTTCAGGGTGTTCAAAGTGGCAGGGAATGGCATCAAGGCATCT	960
QY	329	Glu 329	
DB	961	GAG 963	

14-MAR-2002 (first entry)

Human G-coupled receptor (GCRC) cDNA, Seq ID No 42.

Human; cytostatic; neuroprotective; immunosuppressant; nootropic; anti-inflammatory; anti-viral; gastrointestinal; cardiovascular; cerebroprotective; G-coupled receptor; cell proliferative disease; lymphoma; leukaemia; breast cancer; cirrhosis; neurological disorder; stroke; Alzheimer's disease; multiple sclerosis; mental retardation; cardiovascular disease; atherosclerosis; angina pectoris; indigestion; congestive heart failure; gastrointestinal disorder; dysphagia; AIDS; gastritis; autoimmune disorder; inflammatory disorder; Crohn's disease; systemic lupus erythematosus; metabolic disorder; diabetes; obesity; viral infection; herpesvirus; parvovirus; acquired immune deficiency syndrome; ss.

Homo sapiens.

WO200190359-A2.

29-NOV-2001

22-MAY-2001; 2001WO-US016833.

22-MAY-2000; 2000US-0206222P-

22-MAY-2000; 2000US-020622ZP.
25-MAY-2000; 2000US-0207476P.

23-MAY-2000; 2000US-0207476P.
02-JUN-2000; 2000US-0208834P

02-JUN-2000; 2000US-0208834P.
02-JUN-2000; 2000US-0208861P

02-JUN-2000; 2000US-0208861P.
07-JUN-2000; 2000US-0209868P

(INCY-) INCYTE GENOMICS INC.

Patterson C, Tribouley CM, Yao MG, Griffin JA, Thornton M, Lu Y; Kalliock DA, Candi A, ...

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

WPI; 2002-106199/
P-PSDB: AAU80511

New G-protein coupled receptors useful for treating or preventing cell proliferative (e.g. leukemia), neurological (e.g. stroke), cardiovascular or autoimmune/inflammatory disorders.

Claim 5; Page 146; 148pp; English.

The invention relates to a novel human G-coupled receptor (1). (1) and its corresponding polynucleotides are useful for diagnosing, treating or preventing cell proliferative diseases (e.g. lymphoma, leukaemia, breast cancer or cirrhosis), neurological disorders (e.g. stroke, Alzheimer's disease, multiple sclerosis or mental retardation), cardiovascular diseases (e.g. atherosclerosis, angina pectoris or congestive heart failure), gastrointestinal disorders (e.g. dysphagia, indigestion or gastritis), autoimmune/inflammatory disorders (e.g. AIDS, Crohn's disease) or systemic lupus erythematosus) or metabolic disorders (e.g. diabetes or obesity), or viral infections (e.g. infection by herpesvirus or parvovirus). ABK16615-ABK16637 represent novel human G-coupled receptor coding sequences of the invention.

sequence 966 BP; 182 A; 276 C; 217 G; 291 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	1,348-164	Length:	966
Score:	1649.00	Matches:	317
Percent Similarity:	99.07%	Conservative:	1
Best Local Similarity:	98.75%	Mismatches:	3
Query Match:	95.98%	Indels:	0
DB:	6	Gaps:	0

US-10-081-775-2 (1-329) x ABK16633 (1-966)

9 MetGluSerProHisHisThrAspValAspProSerValPheLeuLeuGlyIlePro 28

RESULT 9
ABK16633
ID ABK1
XX
AC ABK1

[illegible]

Qy	29	GlyLeuGluGlnPheHisLeuTrrIleuSerLeuProValCysGlyLeuGlyThrAlaThr	48
Dd	61	GGFTCTGGAAACAATTTTCATTGTGGCTCTCACTCCCTGTGTGGCTTAGGCACAGCCACA	120
Qy	49	IleValGlyAsnIleThrIleLeuValValAlaThrGluProValLeuHisLysPro	68
Dd	121	ATTGTGGCCAATAATACTATTCTTGTTGTTGTTGCCACTGAACAGTCTTGCAAGAAGCCT	180
Qy	69	ValTyrlleuPheLeuCysMetLeuSerThrIleAaspLeuAlalaSerValSerThrVal	88
Dd	181	GTGTACCTTTTCTGTGCATGCTCTCAACCATCGACTTGGCTGCCTCTGTCTCCACAGTT	240
Qy	89	ProLysLeuLeuAlaIlePheTrpCysGlyAlaGlyHisIleSerAlaSerAlaCysLeu	108
Dd	241	CCCAGAAGCTACTGGCTATCTTCTGTGTGGAGCCGACATATATCTGCTCTGCTGCCTG	300
Qy	109	AlaHisMetPhePheIleHisAlaPheCysMetMetGluSerThrValLeuLeuAlaMet	128
Dd	301	GCACAGATGTTCTTCATTTCATGCTTCTGCATGATGAGTCCACTGTGTACTTGGCCATG	360
Qy	129	AlaPheAspArgTyrValAlaIleCysHisProLeuArgTyrAlaThrIleLeuThrAsp	148
Dd	361	GCCTTTGATCGCTACGTGGCCATCTGCCACCCTCTCCGTATGGCACAAATCCTCACTGAC	420
Qy	149	ThrIleIleAlaHisIleGlyValAlaAlaValValArgGlySerLeuLeuMetLeuPro	168
Dd	421	ACCATCATTTGCCACATAGGGGTGGCAGCTGTAGTGGAGGCTCCCTGTCTCATGCTCCA	480
Qy	169	CysProPhePheIleGlyArgLeuAsnPheCysGlnSerHisValIleLeuHisThrTyr	188
Dd	481	TGTCCCTTTCCTATTGGCGGTTTGAACCTTCTGCCAAAGCCATGTGATCCTACACGTCAC	540
Qy	189	CysGluHisMetAlaValValLysLeuAlaCysGlyAspThrArgProAsnArgValTyr	208
Dd	541	TGTGAGCACATGGCTGTGGTGAAGCTGGCCCTGTGGAGACACCGAGCCCTAACCGTGTAT	600
Qy	209	GlyLeuThrAlaAlaLeuLeuValIleGlyValAspLeuPheCysIleGlyLeuSerTyr	228
Dd	601	GGGCTGCACAGCTGCACTGTGTGTCAATGGGGTGTGATTTGTTTGCATTGGTCTCTCTAT	660
Qy	229	AlaLeuIleAlaGlnAlaValLeuArgLeuSerSerHisGluAlaArgSerLysAlaLeu	248
Dd	661	GGCCTTAAGTGCACAAGCTGTCTTGGCTCTCATCCCATGAAGCTGGTCCGAAGGCCCTTA	720
Qy	249	GlyThrCysGlySerHisValCysValIleLeuIleSerTyrThrProAlaLeuPheSer	268
Dd	721	GGGACCTGGTTCCTCATGTCTGTGTCATCTCTTATACACAGCCCTCTTCTTCC	780
Qy	269	PhePheThrHisArgPheGlyHisHisValProValHisIleHisIleLeuLeuAlaAsn	288
Dd	781	TTTTTTACACCGCTTTGGCCATCAGTTCAGTCCATATTCATTCATCTTTTGGCCAAAT	840
Qy	289	ValTyrlleuLeuLeuProProAlaLeuAnProValValTyrGlyValLysThrLysGln	308
Dd	841	GTTTATCTGCTTTTGGCACCTGCTCTTAACTCTGTGTGTATGTGAGTGTAAAGACCAAACAG	900
Qy	309	IleArgLysArgValValArgValPheGlnSerGlyGlnGlyMetGlyLysAlaSer	328

DB 901 ATCCGTAATAGAGAGTGGTCAGGGTGTTCTCAGAGTGGGACGGCAATCGGCATTGCGCAT

QY	329	Glu 329
Db	961	GAG 963

RESULT 10
ABK68612
ID ABK68612 standard; DNA; 966 BP.
XX
XX AC ABK68612;
XX
XX 02-JUL-2002 (first entry)
DT
XX

QY 29 GlyLeuGluGlnPheHisLeuTrpLeuSerLeuProValCysGlyLeuGlyThrAlaThr 48
 DB 61 GGTCGGAACAATTTTCATTGTGGCTCTACTCCCTGTGTGGCTTAGGCACAGCCACA 120
 QY 49 IleValGlyAsnIleThrLeuValValAlaThrGluProValLeuHisLysPro 68
 DB 121 ATTGTGGCAATAATAACTATCTGTGGTTGTGGTGGCTGAACAGCTCTTGACAGCCT 180
 QY 69 ValTyrLeuPheLeuCysMetLeuSerThrIleAspLeuAlaAlaSerValSerThrVal 88
 DB 181 GTGTACCTTTTCTGTGCAGCTCTCAACCATCGACTTGGCTGTCTCTCCACAGTT 240
 QY 89 ProLysLeuLeuAlaIlePheTrpCysGlyAlaGlyHisIleSerAlaSerAlaCysLeu 108
 DB 241 CCCAAGCTACTGGCTATCTTCTGGTGTGGAGCGGACATATCTGCCCTCTGCCGCTCG 300
 QY 109 AlaHisMetPhePheIleHisAlaPheCysMetMetGluSerThrValLeuAlaMet 128
 DB 301 GCACAGATGTTCTTTCATTTCATGCTTCTGCATGATGGAGTCCACTGTGCTACTGGCCATG 360
 QY 129 AlaPheAspArgTyrValAlaIleCysHisProLeuArgTyrAlaThrIleLeuThrAsp 148
 DB 361 GCCTTTGATCGTACGTGGCCATCTGCCACCACCTCCGCTATGCCACATCTCTCACTGAC 420
 QY 149 ThrIleIleAlaHisIleGlyValAlaAlaValAlaValArgGlySerLeuLeuMetLeuPro 168
 DB 421 ACCATCATTTGCCACATAGGGGTGGCAGCTGTAGTGGAGGCTCCCTGCTCATGCTCCCA 480
 QY 169 CysProPhePheIleGlyArgLeuAsnPheCysGlnSerHisValIleLeuHisThrTyr 188
 DB 481 TGTCCCTTCTTATTGGGGTGTGAACCTTCTGCCAAGCCATGTGATCTTACACAGTAC 540
 QY 189 CysGluHisMetAlaValIleValIleValIleValIleValIleValIleValIleVal 208
 DB 541 TGTGAGCATGCTGCTGCTGAGAGCTGGCCTGTGGAGACACACAGGCTTAACCGTGTAT 600
 QY 209 GlyLeuThrAlaAlaLeuValIleGlyValAspLeuPheCysIleGlyLeuSerTyr 228
 DB 601 GGGCTGACAGCTGACCTGTGTGTCATTTGGGTGACTTGTGTTGCAITGGTCTCTCCAT 660
 QY 229 AlaLeuIleAlaGlnAlaValLeuArgLeuSerSerHisGluAlaArgSerLysAlaLeu 248
 DB 661 GCCCTAAGTGCACAGCTGCTCTCGCCTCTCATCCCATGAAGCTCGTCCCAAGCCCTA 720
 QY 249 GlyThrCysGlySerHisValCysValIleLeuIleSerTyrThrProAlaLeuPheSer 268
 DB 721 GGGACCTGTGGTTCCTCATCTGCTGATCTCTTATATATATATATATATATATATAT 780
 QY 269 PhePheThrHisArgPheGlyHisHisValProValHisIleHisIleLeuLeuAlaAsn 288
 DB 781 TTTTATACACACCGCTTTGGCCATCATGTTCCAGTCCATATTCATCTTTTGGCCAT 840
 QY 289 ValTyrLeuLeuLeuProProAlaLeuAsnProValValTyrGlyValLysThrLysGln 308
 DB 841 GTTATCTGCTTTTGGCCACCTGCTCTTAATCTCTGTGTATATATATATATATATATAT 900
 QY 309 IleArgLysArgValValArgValPheGlnSerClyGlnGlyMetGlyIleLysAlaSer 328
 DB 901 ATCCGTAAGAAGAGTGTTCAGGGTGTTCAGAGTGGGAGGGAATGGGCATCAAGGCATCT 960
 QY 329 Glu 329
 DB 961 GAG 963
 RESULT 11
 ABZ77974
 ID ABZ77974 standard; DNA; 948 BP.
 XX
 AC
 AC
 AC
 DB 17-APR-2003 (first entry)
 XX

DE XX Human G protein coupled receptor coding sequence SEQ ID 205.
 KW Human; gene; anorectic; antiasthmatic; antidiabetic; hypotensive;
 KW antiparkinsonian; nootropic; neuroprotective; tranquiliser;
 KW antirheumatic; antiinflammatory; osteopathic; cardiant; neuroleptic;
 KW antiarthritic; gene therapy; olfactory G protein-coupled receptor; GPCR;
 KW infection; obesity; diabetes; hypertension; malnutrition;
 KW Parkinson's disease; Alzheimer's disease; Korsakoff's psychosis; anxiety;
 KW rheumatoid arthritis; chronic obstructive pulmonary disease;
 KW osteoporosis; asthma; myocardial infarction; schizophrenia;
 KW osteoarthritis; ds.
 XX Homo sapiens.
 OS WO2003000735-A2.
 PN 03-JAN-2003.
 PD 24-JUN-2002; 2002WO-IB002481.
 PF 26-JUN-2001; 2001US-0301095P.
 PR 06-NOV-2001; 2001US-0332758P.
 XX (DECO-) DECODE GENETICS EHF.
 PA Martinez RAM, Sigurdsson GT;
 PI P-PSDB; ABR01673.
 DR WPI; 2003-175284/17.
 DR P-PSDB; ABR01673.
 DR New olfactory G protein-coupled receptor gene nucleic acid and
 PT polypeptide, useful for diagnosing or treating a disease or condition
 PT associated with GPCR, e.g. obesity, diabetes, hypertension, malnutrition
 PT or Alzheimer's disease.
 PS Claim 1; Page 103; 383pp; English.
 CC The present invention relates to novel human olfactory G protein-coupled
 CC receptors (GPCR) and their coding sequences (ABZ77872-ABZ77986 and
 CC ABR01571-ABR01685). The GPCRs and coding sequences are useful for
 CC diagnosing or treating a disease or condition associated with GPCR, e.g.
 CC infections, obesity, diabetes, hypertension, malnutrition, Parkinson's
 CC disease, Alzheimer's disease, Korsakoff's psychosis, anxiety, rheumatoid
 CC arthritis, chronic obstructive pulmonary disease, osteoporosis, asthma,
 CC myocardial infarction, schizophrenia, or osteoarthritis
 XX SQ Sequence 948 BP; 174 A; 266 C; 202 G; 306 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 1.01e-135 Length: 948
 Score: 1376.00 Matches: 267
 Percent Similarity: 94.16% Conservative: 7
 Best Local Similarity: 91.75% Mismatches: 17
 Query Match: 80.09% Indels: 0
 DB: Gaps: 0
 US-10-081-775-2 (1-329) x ABZ77974 (1-948)
 QY 20 SerValPhePheLeuLeuGlyIleProGlyLeuGluGlnPheHisLeuTrpLeuSerLeu 39
 DB 28 TCTGAGTTTATCTGATGGGATTCCTGGCATTCACACTTCTTGTGTTTGTGTTT 87
 QY 40 ProValCysGlyLeuGlyThrAlaThrIleValGlyAsnIleThrIleLeuValValVal 59
 DB 88 TTGTTGTTCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 147
 QY 60 AlaThrGluProValLeuHisLysProValTyrLeuPheLeuCysMetLeuSerThrIle 79
 DB 148 GCCACTGAACCACTCTTGCAAGCCTGTGACCTTTTCTGTGATGCTCTCAACATC 207
 QY 80 AspLeuAlaAlaSerValSerThrValProLysLeuLeuAlaIlePheTrpCysGlyAla 99
 DB 208 GACTTGGCTGCTCTGTCTCCACAGTCTCCAGTCTCTGCTATCTCTGTTGGAGCC 267

QY 100 GlyHisIleSerAlaSerAlaCysLeuAlaHisMetPhePheIleHisAlaPheCysMet 119
 DB 268 GGACATATATCTGCTCTGCTGCGCACAGATGTTCTTCATTCATGCTTCTGCGATG 327
 QY 120 MetGluSerThrValLeuLeuAlaMetAlaPheAspArgTyrValAlaIleCysHisPro 139
 DB 328 ATGGAGTCCACTGTGCTACTGGCCATGGCCCTTTGATCGTACGTGGCCATCTGCCACCCA 387
 QY 140 LeuArgTyrAlaThrIleLeuThrAspThrIleIleAlaHisIleGlyValAlaAlaVal 159
 DB 388 CTCCGCTATGCCCAATCTCTCACTGACACCATCATTTGCCACATAGGGGTGGCAGCTGTA 447
 QY 160 ValArgGlySerLeuLeuMetLeuProCysProPhePheIleGlyArgLeuAsnPhcCys 179
 DB 448 GTGGAGGCTCCCTGCTCATGCTCCCATGTCCCTTCTTATTTGGGCGTTTGAATCTTCGC 507
 QY 180 GlnSerHisValIleLeuHisThrTyrCysGluHisMetAlaValValLysLeuAlaCys 199
 DB 508 CAAAGCCATGTGATCTTACACACGCTACTGTGAGCACATGGCTGTGGTGAAGCTGGCTGT 567
 QY 200 GlyAspThrArgProAsnArgValTyrGlyLeuThrAlaAlaLeuLeuValIleGlyVal 219
 DB 568 GGAGACACCGCTTACCGGTGTATGGGTGACAGCTGCACTGTGTGCTCATTTGGGTT 627
 QY 220 AspLeuPheCysIleGlyLeuSerTyrAlaLeuIleAlaGlnAlaValLeuArgLeuSer 239
 DB 628 GACTTGTGTTGCTATGCTCTCTATGCCCTAAAGTGCACAGCTGCTCTTCGCTCTCA 687
 QY 240 SerHisGluAlaArgSerLysAlaLeuGlyThrCysGlySerHisValCysValIleLeu 259
 DB 688 TCCCATGAAGCTCGGTCCAGGCCCTAGGACCTGTGGTTCCTATGCTGTGCTATCTC 747
 QY 260 IleSerTyrThrProAlaLeuPheSerPhePheThrHisArgPheGlyHisValPro 279
 DB 748 ATCTCTTATACACAGCCCTCTCTCTCTTTTACACACCGCTTTGGCCATCAGCTTCCA 807
 QY 280 ValHisIleHisIleLeuLeuAlaAsnValTyrLeuLeuLeuProProAlaLeuAsnPro 299
 DB 808 GTCCATATTCATATCTTTTGGCCATGTTATCTGCTTTTGGCCACCTGCTCTTAATCCT 867
 QY 300 ValValTyrGlyValLysThrLysGlnIleArg 310
 DB 868 GTGGTATATGAGGTAAAGACCAACAGATCCAA 900
 RESULT 12
 ID ADC86380
 AC ADC86380 standard; DNA; 1300 BP.
 DT 01-JAN-2004 (first entry)
 DE Human GPCR gene SEQ ID NO:833.
 XX ds; gene; human; GPCR;
 KW guanosine triphosphate-binding protein coupled receptor; gene therapy.
 OS Homo sapiens.
 XX EP1270724-A2.
 PN 02-JAN-2003.
 XX 18-JUN-2002; 2002EP-00013517.
 XX 18-JUN-2001; 2001JP-00246789.
 PR (NAAD-) NAT INST ADVANCED IND SCI & TECHNOLOGY.
 XX (ADSC-) CENT ADVANCED SCI & TECHNOLOGY INCUBATIO.
 PA Suwa M, Asai K, Akiyama Y, Aburatani H;
 XX

DR WPI; 2003-315783/31.
 DR P-PSDB; ADC86381.
 XX New polynucleotide, useful for preparing a composition for treating a
 PT patient in need of increased or suppressed activity or expression of the
 PT guanosine triphosphate-binding protein coupled receptor.
 XX Claim 1; SEQ ID NO 833; 28pp; English.
 XX The invention relates to a novel polynucleotide encoding a guanosine
 CC triphosphate-binding protein coupled receptor (GPCR). A polynucleotide of
 CC the invention may have a use in gene therapy. The polynucleotide and
 CC polypeptide are useful for preparing a composition for treating a patient
 CC in need of increased or suppressed activity or expression of the
 CC guanosine triphosphate-binding protein coupled receptor. The
 CC polynucleotide sequences shown in ADC85548-ADC87616 encode GPCR's of the
 CC invention.
 XX Sequence 1300 BP; 274 A; 379 C; 299 G; 348 T; 0 U; 0 Other;
 SQ
 Alignment Scores:
 Pred. No.: 1,81e-93 Length: 1300
 Score: 978.50 Matches: 181
 Percent Similarity: 73.44% Conservatives: 54
 Best Local Similarity: 56.56% Mismatches: 82
 Query Match: 56.96% Indels: 3
 Gaps: 2
 DB:
 US-10-081-775-2 (1-329) x ADC86380 (1-1300)
 QY 1 MetSerSerThrLeuGlyHisAsnMetGluSerProHisHisHisThrAspValAspProSer 20
 DB 135 CTAAACACCTTACTTGGCCATGAAACTCATAAACCATACCATCAG-----AACCCAAACC 188
 QY 21 ValPhePheLeuLeuGlyIleProGlyLeuGluGlnPheHisLeuTyrLeuSerLeuPro 40
 DB 189 TCCCTTTCCTCATGGGAATTCAGGCCCGGAGGAGCATCCACTTTTGGATGCTTTCCC 248
 QY 41 ValCysGlyLeuGlyThrAlaThrIleValGlyAsnIleThrIleLeuValValAla 60
 DB 249 TTCCTCTCATGTATGCCCTGGCAGTGTGGGAAACATGGTGGTCTGTAGTGTATCAT 308
 QY 61 ThrGluProValLeuHisLysProValTyrLeuPheLeuCysMetLeuSerThrIleAsp 80
 DB 309 TCAGAGCCTGTATTGCACAGCCCATGTACCTTCTCTCATGCTATCCACCATGAC 368
 QY 81 LeuAlaAlaSerValSerThrValProLysLeuLeuAlaIlePheTyrCysGlyAlaGly 100
 DB 369 CTGGTCTCTGCACCTCCACTGTGCCCAAGCTCTTGCATCTTATCCATGGCTTCTCAGCTGTA 428
 QY 101 HisIleSerAlaSerAlaCysLeuAlaHisMetPhePheIleHisAlaPheCysMetMet 120
 DB 429 GAGATCAACTTTGGGCGCTGTCTGCCAGATGTTCTTTATCCATGGCTTCTCAGCTGTA 488
 QY 121 GluSerThrValLeuLeuAlaMetAlaPheAspArgTyrValAlaIleCysHisProLeu 140
 DB 489 GAATCTGTATATCTGTAGCAATGGCTTTGACCGCTACTTAGCCATTTTGGGCGCTCTG 548
 QY 141 ArgTyrAlaThrIleLeuThrAspThrIleIleAlaHisIleGlyValAlaAlaValVal 160
 DB 549 CACTATGGGTCAITGTCTCTCCCGAGTCTGTAGGCAAGCTGGGGCTGCAGCCGTGCTT 608
 QY 161 ArgGlySerLeuLeuMetLeuProCysProPhePheIleGlyArgLeuAsnPhcCysGln 180
 DB 609 CGTGGTTTGGGACTCATGCCCTCACTCACCTGCTTATGGCAAGCTAGACTACTGC--- 665
 QY 181 SerHisValIleLeuHisThrTyrCysGluHisMetAlaValValLysLeuAlaCysGly 200
 DB 666 AGTCGAGTGGTGGCCCACTCTACTGTGAACATGCTGTGTGTAAAGCTTGGCTTTGGA 725
 QY 201 AspThrArgProAsnArgValTyrGlyLeuThrAlaAlaLeuLeuValIleGlyValAsp 220
 DB 726 GGAACACAGCCCAACAAACATCATGTATGTCATCCTGCTGCCACACTGTGTGGTGGGCACTGAC 785

PD 19-APR-2001.
 XX 06-OCT-2000; 2000WO-US027582.
 XX 08-OCT-1999; 99US-015861SP.
 PR 24-FEB-2000; 2000US-0184809P.
 XX (DIGI-) DIGISCENTS.
 PA (YEDA) YEDA RES & DEV CO LTD.
 XX
 XX Ballenson J, Smith D, Lancet D, Glusman G, Fuchs T, Yanai I;
 XX WPI; 2001-290713/30.
 XX
 PT New polynucleotides which encode polypeptides involved in olfactory
 PT sensation for identifying olfactory agonists and antagonists.
 XX
 PS Claim 8; Page 431; 1857pp; English.
 CC
 CC The present sequence is one of a number of isolated polynucleotides which
 CC encode polypeptides involved in olfactory sensation. The polynucleotides
 CC can be used in screening for olfactory agonists and antagonists. The
 CC methods allow for the determination of primary scents and the
 CC identification of the odour receptors used to detect these primary
 CC scents. The methods also enable determination of secondary scents and the
 CC identification of combinations of odour receptors that are involved in
 CC detecting such secondary scents. This enables the construction of a scent
 CC representation (also called a scent fingerprint or scent profile), which
 CC may be used to re-create and edit scents. Libraries of olfactory
 CC receptors are useful for determining the interaction pattern of a
 CC composition with the receptors, and can be used for determining
 CC differences in the olfactory faculties of different individuals
 XX
 SQ Sequence 985 BP; 193 A; 300 C; 225 G; 267 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 1.15e-88 Length: 985
 Score: 931.50 Matches: 177
 Percent Similarity: 73.79% Conservative: 51
 Best Local Similarity: 57.28% Mismatches: 78
 Query Match: 54.22% Indels: 4
 DB: 4 Gaps: 2

US-10-081-775-2 (1-329) x AAH32052 (1-985)

QY 12 ProHisHisThrAspValAspProSerValPhePheLeuLeuGlyIleProGlyLeuGlu 31
 DB 15 CCATACCATCAG-----AACCCACCTCCCTTCTGCTCATGGGAATCCAGGCCCGAG 68
 QY 32 GlnPheHisLeuTrpLeuSerLeuProValCysGlyLeuGlyThrAlaThrIleValGly 51
 DB 69 GCATCCCATTTTGGATTGGCTTTTCCCTCTGCTCCATGATGCCCTGGGACCTGGGA 128
 QY 52 AsnIleThrIleLeuValValAlaThrGluProValLeuHisHisProValTyrLeu 71
 DB 129 AACATGGTGGTGGCTAGTGGTATCAGAGCCTGATTGACACCCATGACCTG 188
 QY 72 PheLeuCysMetLeuSerThrIleAspLeuAlaAlaSerValSerThrValProLysLeu 91
 DB 189 TTCTCTGCATGATATCCACATTTGACCTGGTCTGCTGCACCTCCACTGTGCCAAGCTC 248
 QY 92 LeuAlaIlePheTrpCysGlyAlaGlyHisIleSerAlaSerAlaCysLeuAlaHisMet 111
 DB 249 CTTGCACCTTTTGGGCAAGGATGCTGAGATCAACTTTGGGGCTGTGCTGCCAGAT 308
 QY 112 PhePheIleHisAlaPheCysMetMetGluSerThrValLeuLeuAlaMetAlaPheAsp 131
 DB 309 TTTCTTATCATCGCTTCTCAGCTAGATAATCTGGTATCTGTAGCAATGGCTTTGAC 368
 QY 132 ArgTyrValAlaIleCysHisProLeuArgTyrAlaThrIleLeuThrAspThrIleIle 151
 DB 369 CGCTACTTAGCCATTTGCTGGCCTCTGCACTATGGGTCTATGCTCTCCCGAGAGTCTGTA 428

QY 152 AlaHisIleGlyValAlaAlaValValArgGlySerLeuLeuMetLeuProCysProPhe 171
 DB 429 GGCAAGCTGGGGGCTGCAGC-GTGCCTCGTGGTTTGGGACTCATGACCCCACTCCTGC 487
 QY 172 PheIleGlyArgLeuAsnPheCysGlnSerHisValIleLeuHisThrTyrCysGluHis 191
 DB 488 TTACTGGCAAGACTGAGCTACTGC--AGTCGAGTGGTGGCCCACTCCTACTGTGAACAC 544
 QY 192 MetAlaValValLysLeuAlaCysGlyAspThrArgProAsnArgValTyrGlyLeuThr 211
 DB 545 ATGGCTGTGGTAAAGCTGGCTTGTGGAGGAACACAGCCAAACACATCTATGGCATCACT 604
 QY 212 AlaAlaLeuLeuValIleGlyValAspLeuPheCysIleGlyLeuSerTyrAlaLeuIle 231
 DB 605 GCTGCCACACTGGTGGTGGGCACTGACTCCATCTGTATTGCTCTCTTATGCACTCAIC 664
 QY 232 AlaGlnAlaValLeuArgLeuSerHisGluAlaArgSerLysAlaLeuGlyThrCys 251
 DB 665 CTCGGAGCTGTGTAGGCTTTTCTCCAGAGGCAAGGGCTAAGACCTTTGGCACTTGT 724
 QY 252 GlySerHisValCysValIleLeuIleSerTyrThrProAlaLeuPheSerPheThr 271
 DB 725 GGCTCCCACTGGGTGTATATCTTCTTCTACACCAAGGACTCTCTCTCTTCTACACA 784
 QY 272 HisArgPheGlyHisHisValProValHisIleHisIleLeuLeuAlaAlaValTyrLeu 291
 DB 785 CAGCGGTTTGGCCAGCAGCGTGCCTGGCAGCATCCACATCTTCTAGCTGACCTTACCTG 844
 QY 292 LeuLeuProProAlaLeuAsnProValValTyrGlyValLysThrLysGlnIleArgLys 311
 DB 845 GTTGTGCCACCCCATGCTCAACCCCATCATCTATGCGATGAAGACCAACAGATCTGGGAT 904
 QY 312 ArgValValArgValPheGlnSerGly 320
 DB 905 GGGGCCCTCCGGCTTCTGAAGTGGGGC 931

Search completed: August 27, 2004, 18:47:31
 Job time : 483 secs

ms Page Book (uspio)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 27, 2004, 14:41:51 ; Search time 478 Seconds

(without alignments)
8771.907 Million cell updates/sec

Title: US-10-081-775-1_COPY_537_1523

Perfect score: 987

Sequence: 1 atgtccagcactcttgccca.....tgggcatcaagcattctgag 987

Scoring table: IDENTITY_NUC

; Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 212409041 residues

Total number of hits satisfying chosen parameters: 6747725

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_29Jan04:*

- 1: Geneseqn1980s:*
- 2: Geneseqn1990s:*
- 3: Geneseqn2000s:*
- 4: Geneseqn2001as:*
- 5: Geneseqn2001bs:*
- 6: Geneseqn2002as:*
- 7: Geneseqn2003as:*
- 8: Geneseqn2003bs:*
- 9: Geneseqn2003cs:*
- 10: Geneseqn2004as:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	987	100.0	1567	8	ACD91434 Human G-p
2	980.6	99.4	990	5	AAS42339 Human cdn
3	980.6	99.4	990	6	ABZ43046 Human GPC
4	980.6	99.4	990	6	ABK37625 DNA encod
5	980.6	99.4	998	6	ABQ88371 Human G p
6	980.6	99.4	998	6	ABQ88372 Human G p
7	980.6	99.4	1390	9	ADC86344 Human GPC
8	956.6	96.9	963	4	Aah31850 Human olf
9	956.6	96.9	966	6	ABK16633 Human G-c
10	956.6	96.9	966	6	ABK68612 Human G-c
11	791	80.1	948	7	ABZ77974 Human DNA
12	376.2	38.1	980	4	AAD19143 Human G p
13	376.2	38.1	1986	4	AAD113897 Human G-p
14	376.2	38.1	1986	4	ABAS5607 Human G-p
15	376.2	38.1	1986	4	ABAS5607 Human G-p
16	376.2	38.1	1986	4	AAI35262 Human G-p
17	376.2	38.1	1986	4	ABA45127 Human G-p
18	376.2	38.1	1986	4	ABA25305 Human G-p
19	376.2	38.1	1986	4	AAK29302 Human G-p
20	376.2	38.1	1986	4	AAK03835 Human G-p
21	376.2	38.1	1986	4	ABZ8924 Human G-p
22	376.2	38.1	1986	5	AAI03764 Human G-p
23	376.2	38.1	1986	6	ABZ03861 Human G-p
24	376.2	38.0	942	4	Aah32015 Human Olf

24	375.2	38.0	945	4	AAD12947	Aad12947 Human G-p
25	375.2	38.0	945	5	AAS42253	Aas42253 Human cdn
26	375.2	38.0	945	6	AAL44139	Aal44139 Human G-p
27	375.2	38.0	945	6	ABZ42950	Abz42950 Human GPC
28	375.2	38.0	945	6	ABK37539	Abk37539 DNA encod
29	374.6	38.0	1345	9	ADC86304	Adc86304 Human GPC
30	373.6	37.9	945	6	ABK68589	Abk68589 Human DNA
31	369.8	37.5	945	4	AAI23114	Aai23114 Probe #13
32	369.8	37.5	945	4	ABA68205	Abas68205 Human G-p
33	369.8	37.5	945	4	AAI48420	Aai48420 Probe #17
34	369.8	37.5	945	4	ABA50263	Abas50263 Human G-p
35	369.8	37.5	945	4	ABA35217	Abas35217 Probe #13
36	369.8	37.5	945	4	AAK42341	Aak42341 Human G-p
37	369.8	37.5	945	4	AAK16586	Aak16586 Human G-p
38	369.8	37.5	945	4	ABSA1952	Absa1952 Human G-p
39	369.8	37.5	945	5	AAI08767	Aai08767 Probe #87
40	369.8	37.5	945	6	ABS16397	Abs16397 Human G-p
41	368.4	37.3	1013	6	ABQ88366	Abq88366 Human G-p
42	368.4	37.3	1360	9	ADC86368	Adc86368 Human GPC
43	367.2	37.2	982	4	AAD19141	Aad19141 Human G-p
44	365	37.0	1345	9	ADC86302	Adc86302 Human GPC
45	364	36.9	942	4	AAH32013	Aah32013 Human Olf

ALIGNMENTS

RESULT 1

ACD91434	ACD91434 standard; cdna; 1567 BP.
XX	XX
AC	ACD91434;
XX	XX
DT	22-SEP-2003 (first entry)
XX	XX
DE	Human G-protein coupled receptor HGPRMY25 cdna.
XX	XX
KW	Human; G-protein coupled receptor; HGPRMY25; immune disease;
KW	inflammatory disease; arthritis; asthma; AIDS; psoriasis;
KW	graft-versus-host disease; systemic lupus erythematosus;
KW	reproductive disorder; varicocele; orchitis; neural disorder;
KW	Alzheimer's disease; Parkinson's disease; depression; schizophrenia;
KW	cardiovascular disorder; hypertension; acute heart failure;
KW	pulmonary disorder; endocrine disorder; obesity; diabetes; anorexia;
KW	bone disorder; osteoporosis; pain; cancer; chromosome identification;
KW	gene therapy; gene; ss.
XX	XX
OS	Homo sapiens.
XX	XX
PN	US2003060409-A1.
XX	XX
PD	27-MAR-2003.
XX	XX
PF	21-FEB-2002; 2002US-00081775.
XX	XX
PR	21-FEB-2001; 2001US-0270134P.
XX	XX
PR	27-MAR-2001; 2001US-0278952P.
XX	XX
PA	(RAMA/) RAMANATHAN C S.
XX	XX
PA	(FEDE/) FEDER J N.
XX	XX
PA	(MINT/) MINTIER G A.
XX	XX
PI	Ramanathan CS, Feder JN, Mintier GA;
XX	XX
DR	WPI; 2003-521919/49.
XX	XX
DR	P-PSDB; ABO42809.
XX	XX
PT	New nucleic acid molecule encoding a human G-protein coupled receptor
XX	XX
PT	(HGPRMY25) is useful for diagnosing, preventing or treating diseases
XX	XX
PT	involving the receptor, e.g. inflammation, diabetes, asthma, hypertension
XX	XX
PS	or cancer.
XX	XX
PS	Claim 1; Fig 1A-B; 139pp; English.

XX The invention describes an isolated nucleic acid molecule comprising a
CC sequence that is at least 95% identical to a polynucleotide encoding
CC novel human G-protein coupled receptor HGPBMV25. The nucleic acid
CC molecule, polypeptide and antibody are useful in diagnosing, preventing,
CC treating or ameliorating medical conditions where GPCR is directly or
CC indirectly involved, such as immune or inflammatory diseases (e.g.
CC arthritis, asthma, AIDS, graft-versus-host disease, psoriasis or systemic
CC lupus erythematosus), reproductive disorders (e.g. varicocele or
CC orchitis), neural disorders (e.g. Alzheimer's disease, Parkinson's
CC disease, depression or schizophrenia), cardiovascular disorders (e.g.
CC hypertension or acute heart failure), pulmonary disorders, endocrine
CC disorders (e.g. obesity, diabetes or anorexia), bone disorders (e.g.
CC osteoporosis), pain or cancer. The polynucleotide may also be used in
CC chromosome identification, in identifying organisms from minute
CC biological samples, or as molecular weight markers. This sequence encodes
CC a novel human G-protein coupled receptor HGPBMV25
XX
SQ Sequence 1567 BP; 367 A; 382 C; 314 G; 504 T; 0 U; 0 Other;
Query Match 100.0%; Score 987; DB 8; Length 1567;
Best Local Similarity 100.0%; Pred. No. 7e-310;
Matches 987; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGTCAGACATCTTGGCCACAACATGGAATCTCTCATCACATGATGTGACCTTCT 60
DB 537 ATGTCAGACATCTTGGCCACAACATGGAATCTCTCATCACATGATGTGACCTTCT 596
QY 61 GTCCTTCTCCCTGGGATCCAGTCTGGAAACAATTTCAATTTGGTGTCTCACTCCCT 120
DB 597 GTCCTTCTCCCTGGGATCCAGTCTGGAAACAATTTCAATTTGGTGTCTCACTCCCT 656
QY 121 GTGTGTGGCTTAGGCACAGCCACAATTTGGGCAATATAAATTAATTTCTGTTGTGCC 180
DB 657 GTGTGTGGCTTAGGCACAGCCACAATTTGGGCAATATAAATTAATTTCTGTTGTGCC 716
QY 181 ACTGAACAGTCTTGACAGAGCTGTGACCTTTTCTGTGCATGCTCTCAACCATCGAC 240
DB 717 ACTGAACAGTCTTGACAGAGCTGTGACCTTTTCTGTGCATGCTCTCAACCATCGAC 776
QY 241 TTGGCTGCTCTGTCTCCACAGTCTCCAGCTACTGGCTATCTTCTGTGTGGAGCCGA 300
DB 777 TTGGCTGCTCTGTCTCCACAGTCTCCAGCTACTGGCTATCTTCTGTGTGGAGCCGA 836
QY 301 CATATATCTGCTCTGCTGCTGGCAATATGTTCTTCATTCATGCTTCTGCAATGATG 360
DB 837 CATATATCTGCTCTGCTGCTGGCAATATGTTCTTCATTCATGCTTCTGCAATGATG 896
QY 361 GAGTCCACTGTCTACTGCGCATGGCTTTGATCGCTAGTGGCCATCGCCACCCACTC 420
DB 897 GAGTCCACTGTCTACTGCGCATGGCTTTGATCGCTAGTGGCCATCGCCACCCACTC 956
QY 421 CGCTATGCCACAATCTCTACTGACACCATCATTTGCCCAATAGGGGTGGCAGCTGATG 480
DB 957 CGCTATGCCACAATCTCTACTGACACCATCATTTGCCCAATAGGGGTGGCAGCTGATG 1016
QY 481 CGAGGCTCCCTGCTGATGCTCCATGCTCCCTTTTATTTGGGGTTTGAACCTTCTCCAA 540
DB 1017 CGAGGCTCCCTGCTGATGCTCCATGCTCCCTTTTATTTGGGGTTTGAACCTTCTCCAA 1076
QY 541 AGCCATGTGATCTACACAGTACTGTGAGCAGCATGGCTGTGTGAAGCTGGCTGTGGA 600
DB 1077 AGCCATGTGATCTACACAGTACTGTGAGCAGCATGGCTGTGTGAAGCTGGCTGTGGA 1136
QY 601 GACACAGGCTTAACCGTGTGTATGGGCTGACAGCTGCACTGTGTGTCATTTGGGGTTGAC 660
DB 1137 GACACAGGCTTAACCGTGTGTATGGGCTGACAGCTGCACTGTGTGTCATTTGGGGTTGAC 1196
QY 661 TTGTTTGTGATTTGGTCTCTCTATGCTTATGCTTATTTGCAAGCTGCTTCTGCTTCATCC 720
DB 1197 TTGTTTGTGATTTGGTCTCTCTATGCTTATTTGCAAGCTGCTTCTGCTTCATCC 1256
QY 721 CATGAAGCTCGGTCCAAGGCCCTAGGACCTGTGGTTCCCATGCTGTGTGTCATCTCTATC 780

DB 1257 CATGAAGCTCGGTCCAAGGCCCTAGGACCTGTGTGTTCCCATGCTGTGTGTCATCTCATC 1316
QY 781 TCTTATACACAGCCCT 840
DB 1317 TCTTATACACAGCCCT 1376
QY 841 CATATTCACATCTCTTTTGGCCCAATGTTATCTGCTTTTGGCCACCTGCTCTTAATCTCTGTG 900
DB 1377 CATATTCACATCTCTTTTGGCCCAATGTTATCTGCTTTTGGCCACCTGCTCTTAATCTCTGTG 1436
QY 901 GTATATGAGTAAAGACCAACACAGATCCGTAAGAGAGTGTGACGGTGTTCAGAGTGTTCAGAGTGGG 960
DB 1437 GTATATGAGTAAAGACCAACACAGATCCGTAAGAGAGTGTGACGGTGTTCAGAGTGGG 1496
QY 961 CAGGAATGGCATCAAGGCATCTGAG 987
DB 1497 CAGGAATGGCATCAAGGCATCTGAG 1523
RESULT 2
AAS42339
ID AAS42339 standard; cDNA; 990 BP.
XX
XX AAS42339;
XX
XX 18-DEC-2001 (first entry)
XX
XX Human cDNA encoding olfactory receptor AOLFRL41.
XX
XX Human; olfactory receptor; G protein-coupled receptor; GPCR; odourant;
KW ss; food additive; cosmetic; fragrance; pharmaceutical additive.
XX
XX Homo sapiens.
XX
XX WO200168805-A2.
XX
XX 20-SEP-2001.
XX
XX 13-MAR-2001; 2001WO-US007771.
XX
XX 13-MAR-2000; 2000US-0188914P.
PR 24-MAR-2000; 2000US-0192033P.
PR 12-APR-2000; 2000US-0198474P.
PR 24-APR-2000; 2000US-0199335P.
PR 26-MAY-2000; 2000US-0207702P.
PR 23-JUN-2000; 2000US-0213849P.
PR 16-AUG-2000; 2000US-0226534P.
PR 07-SEP-2000; 2000US-0230732P.
PR 07-FEB-2001; 2001US-0266862P.
XX
XX (SENO-) SENOMYX INC.
XX
XX Zozulya S;
XX
XX WPI; 2001-570867/64.
DR P-PSDB; AAU24646.
XX
XX Nucleic acids encoding human olfactory G protein-coupled receptors, where
PT useful for screening for compounds involved in olfactory sensation, where
PT the compounds can be used in the food, pharmaceutical and cosmetic
XX industries to customize odors.
XX
XX Claim 1; Page 142; 319pp; English.
XX
XX The invention relates to nucleic acids encoding human olfactory
CC receptors, OR, (a G protein-coupled receptor, GPCR). The OR's
CC specifically recognise molecules, odourants, that elicit specific
CC olfactory sensation. The human olfactory receptors and polynucleotides
CC encoding them are useful for screening a library of chemical compounds
CC for compounds that are involved in olfactory sensation. Modulators of
CC their activity are useful for pharmacological and genetic modulation of
CC olfactory signalling pathways. Therefore, they can be used in the food,
CC

CC Pharmaceutical and cosmetic industries to customise odours and
CC fragrances. The present sequence encodes a human olfactory receptor of
CC the invention
XX
SQ Sequence 990 BP; 188 A; 285 C; 221 G; 296 T; 0 U; 0 Other;

Query Match 99.4%; Score 980.6; DB 5; Length 990;
Best Local Similarity 99.6%; Pred. No. 6.5e-308;
Matches 983; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 1 ATGTCCAGCACTCTTGGCCACAACTGGAATCTCTCATCACACTGATGTTGACCCCTTCT 60
Db 1 ATGTCCAGCACTCTTGGCCACAACTGGAATCTCTCATCACACTGATGTTGACCCCTTCT 60
QY 61 GTCTTCTCTCTCTGGGATCCAGCTCTGGACATTTTCATTTGGGCTCTCATCCCT 120
Db 61 GTCTTCTCTCTCTGGGATCCAGCTCTGGACATTTTCATTTGGGCTCTCATCCCT 120
QY 121 GTGTGTGGCTTAGGCACAGCCCAATTTGTGGCAATATAACTATTCTGTTGTGTGGC 180
Db 121 GTGTGTGGCTTAGGCACAGCCCAATTTGTGGCAATATAACTATTCTGTTGTGTGGC 180
QY 181 ACTGAACCACTTTGGACAGAGCTGTGTACCTTTTCTGTGCACTCTCAACCACTGAC 240
Db 181 ACTGAACCACTTTGGACAGAGCTGTGTACCTTTTCTGTGCACTCTCAACCACTGAC 240
QY 241 TTGGCTGCTCTGTCTCCACAGTCTCCAGCTACTGGCTATCTTCTGTGGTGGAGCCGA 300
Db 241 TTGGCTGCTCTGTCTCCACAGTCTCCAGCTACTGGCTATCTTCTGTGGTGGAGCCGA 300
QY 301 CATATATCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 360
Db 301 CATATATCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 360
QY 361 GAGTCCACTGTCTACTGGCCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420
Db 361 GAGTCCACTGTCTACTGGCCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420
QY 421 CGCTATGCAATCTCTCACTGACACCATCTTGTATGCTGCTGCTGCTGCTGCTGCTGCT 480
Db 421 CGCTATGCAATCTCTCACTGACACCATCTTGTATGCTGCTGCTGCTGCTGCTGCTGCT 480
QY 481 CGAGGCTCCCTGCTCATCTCCATGCTCCCTCTCTTATTGGGCGTTTGAATCTTGCCTAA 540
Db 481 CGAGGCTCCCTGCTCATCTCCATGCTCCCTCTCTTATTGGGCGTTTGAATCTTGCCTAA 540
QY 541 AGCATGTGATCTTACACAGCTACTGTGAGCAGATGGCTGTGGTGAAGCTGGCTGTGGA 600
Db 541 AGCATGTGATCTTACACAGCTACTGTGAGCAGATGGCTGTGGTGAAGCTGGCTGTGGA 600
QY 601 GACACAGGCTTAACCGTGTATGGGCTGACAGCTGACAGCTGTTGGTCAATTTGGGTTGAC 660
Db 601 GACACAGGCTTAACCGTGTATGGGCTGACAGCTGACAGCTGTTGGTCAATTTGGGTTGAC 660
QY 661 TTGTTTTCATTTGCTCTCTCTATGCTCTCTATGCTCTCTATGCTCTCTATGCTCTCTAT 720
Db 661 TTGTTTTCATTTGCTCTCTCTATGCTCTCTATGCTCTCTATGCTCTCTATGCTCTCTAT 720
QY 721 CATGAAGCTGGTCAAGGCTTATGGGCTAGGACCTGTGGTCCCATGCTGTGCTCATCTCATC 780
Db 721 CATGAAGCTGGTCAAGGCTTATGGGCTAGGACCTGTGGTCCCATGCTGTGCTCATCTCATC 780
QY 781 TCTTATACACAGCCCT 840
Db 781 TCTTATACACAGCCCT 840
QY 841 CATATTACATTTCTTTTGGCAATGTTTATCTGCTTTTGGCACTGCTCTTAACTCTGTG 900
Db 841 CATATTACATTTCTTTTGGCAATGTTTATCTGCTTTTGGCACTGCTCTTAACTCTGTG 900
QY 901 GTATATGGAGTTAAGACCAACAGATCCGTAAGAGTGTGTCAAGGTTGTTCAAGTGGG 960
Db 901 GTATATGGAGTTAAGACCAACAGATCCGTAAGAGTGTGTCAAGGTTGTTCAAGTGGG 960

QY 961 CAGGGAATGGCATCAAGGCATCTGAG 987
Db 961 CAGGGAATGGCATCAAGGCATCTGAG 987

RESULT 3
ABZ43046
ID ABZ43046 standard; DNA; 990 BP.
XX ABZ43046;
XX 06-MAR-2003 (first entry)
XX Human GPCR polynucleotide SEQ ID NO 353.
DE Human; GPCR; G protein coupled receptor; signal transduction; olfactory;
KW drug development; gustatory; taste; fragrance; gene; ds.
XX Homo sapiens.
OS
PN WO200216548-A2.
XX
PD 28-FEB-2002.
XX
PF 30-JUL-2001; 2001WO-IB001446.
XX
PR 04-AUG-2000; 2000JP-00237818.
PR 13-FEB-2001; 2001JP-00034434.
XX
PA (NISC-) JAPAN SCI & TECHNOLOGY CORP.
XX
PI Haga T, Takeda S, Mitaku S;
XX
DR WPI; 2002-304118/34.
DR P-P8DB; ABP95772.
XX
Database global search for G protein-coupled receptors, proteins and
PT encoded genes for studying in vivo signal transduction mechanism and
PT identifying targets for drug development.
XX
PS Claim 9; SEQ ID NO 353; 97pp + Sequence Listing; Japanese.

The invention relates to a method for screening G protein-coupled
receptor (GPCR) genes (ABZ42870-ABZ43216) and/or GPCR proteins (ABP95596-
ABP95942) by extracting open-reading frames containing 6-8 transmembrane
domains with 250-1000 amino acid residues to give a gene homologous with
a known GPCR gene. The receptor proteins and encoded genes are useful for
studying in vivo signal transduction mechanism and identifying targets
for drug development e.g. based on olfactory and gustatory receptors in
form of agonists and antagonists by screening intrinsic and extrinsic
ligands as bitter taste inhibitors, taste enhancers and fragrance
improvers. Note: The sequence data for this patent did not form part of
the printed specification, but was obtained in electronic format directly
from WIPO at ftp.wipo.int/pub/published_pct_sequences

Sequence 990 BP; 188 A; 285 C; 221 G; 296 T; 0 U; 0 Other;

Query Match 99.4%; Score 980.6; DB 6; Length 990;
Best Local Similarity 99.6%; Pred. No. 6.5e-308;
Matches 983; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 ATGTCCAGCACTCTTGGCCACAACTGGAATCTCTCATCACACTGATGTTGACCCCTTCT 60
Db 1 ATGTCCAGCACTCTTGGCCACAACTGGAATCTCTCATCACACTGATGTTGACCCCTTCT 60
QY 61 GTCTTCTCTCTCTGGGATCCAGCTCTGGACAAATTTTCATTTGGGCTCTCATCCCT 120
Db 61 GTCTTCTCTCTCTGGGATCCAGCTCTGGACAAATTTTCATTTGGGCTCTCATCCCT 120
QY 121 GTGTGTGGCTTAGGCACAGCCCAATTTGTGGCAATATAACTATTCTGTTGTGTGGC 180
Db 121 GTGTGTGGCTTAGGCACAGCCCAATTTGTGGCAATATAACTATTCTGTTGTGTGGC 180

Db 421 CGCTATGCCCAATCCTCACTGACACACCATCATTCGCCACATAGGGGTGGCGAGCTGTAGTG 480

Qy 481 CGAGGCTCCCTGCTCAGTCTCCCATGTCCCTTCTTTATTGGGCGTTTGAATCTTGGCCAA 540

Db 481 CGAGGCTCCCTGCTCAGTCTCCCATGTCCCTTCTTTATTGGGCGTTTGAATCTTGGCCAA 540

Qy 541 AGCATGTGATCCTACACACAGTACTGTGTGAGCAGATGGCTGTGGTGAAGCTGGCCCTGTGG 600

Db 541 AGCATGTGATCCTACACACAGTACTGTGTGAGCAGATGGCTGTGGTGAAGCTGGCCCTGTGG 600

Qy 601 GACACAGGCTTAACCGTGTGTATGGCTGACAGCTGACACTGTTGGTCAATGGGTTGCAC 660

Db 601 GACACAGGCTTAACCGTGTGTATGGCTGACAGCTGACACTGTTGGTCAATGGGTTGCAC 660

Qy 661 TTGTTTTCATGTGCTCTCCTATGCTCCCTTAATGACAGCTGTCTTCCGCTCTTCATCC 720

Db 661 TTGTTTTCATGTGCTCTCCTATGCTCCCTTAATGACAGCTGTCTTCCGCTCTTCATCC 720

Qy 721 CATGAAGCTCGGTCCAGGCCCTAGGACCTGTGGTTCCTATGTTCCATGTCTGTGTATCCTCATC 780

Db 721 CATGAAGCTCGGTCCAGGCCCTAGGACCTGTGGTTCCTATGTTCCATGTCTGTGTATCCTCATC 780

Qy 781 TCTTATACACAGCCCTCTTCTCTCTTTTACACAGCTGTCTTCCGCTCTTCATCC 840

Db 781 TCTTATACACAGCCCTCTTCTCTCTTTTACACAGCTGTCTTCCGCTCTTCATCC 840

Qy 841 CATATTCATCTTTTGGCCAAATGTTATCTGCTTTTGGCCACCTCTCTTAACTCTGTG 900

Db 841 CATATTCATCTTTTGGCCAAATGTTATCTGCTTTTGGCCACCTCTCTTAACTCTGTG 900

Qy 901 GTATATGGAGTTAAGACCAACAGATCCGTAAAGAGTTGTCAAGGTGTTTCAAGTGGG 960

Db 901 GTATATGGAGTTAAGACCAACAGATCCGTAAAGAGTTGTCAAGGTGTTTCAAGTGGG 960

Qy 961 CAGGGAATGGGCAATCAAGGCATCTGAG 987

Db 961 CAGGGAATGGGCAATCAAGGCATCTGAG 987

RESULT 5

ABQ88371

ID ABQ88371 standard; cDNA; 998 BP.

AC ABQ88371;

XX

XX

XX 23-SEP-2002 (first entry)

XX

XX Human G protein coupled receptor cDNA SEQ ID NO:35.

XX Human; G protein coupled receptor; GPCR; GPCR; neuroprotective;

XX neurotropic; anti-HIV; antiasthmatic; antiarteriosclerotic; cytosstatic;

XX immunomodulator; antiinflammatory; antidiabetic; anorectic; haemostatic;

XX antibacterial; fungicide; protozoal; virucide; nephrotropic; antiparkinsonian; HIV;

XX cardiant; antiulcer; antiallergic; hepatotropic; antiparkinsonian; HIV;

XX vaccine; gene therapy; cell signal processing; cardiomyopathy; diabetes;

XX metabolic pathway modulation; atherosclerosis; cancer; obesity; asthma;

XX infection; Parkinson's disease; osteoporosis; Crohn's disease; ulcer;

XX allergy; cirrhosis; glomerulonephritis; stroke; haematopoietic disorder;

XX systemic lupus erythematosus; gene; ss.

XX

XX Homo sapiens.

XX

XX WO200250276-A2.

XX

XX 27-JUN-2002.

XX

XX 18-DEC-2001; 2001WO-US049347.

XX

XX 18-DEC-2000; 2000US-0256635P.

XX

XX 21-DEC-2000; 2000US-0257878P.

XX

XX 04-JAN-2001; 2001US-0259743P.

XX

XX 10-JAN-2001; 2001US-0260718P.

XX

XX 12-JAN-2001; 2001US-0261498P.

PR 24-JAN-2001; 2001US-0263689P.

PR 08-FEB-2001; 2001US-0267464P.

PR 22-FEB-2001; 2001US-0271021P.

PR 14-MAR-2001; 2001US-0275946P.

PR 23-MAR-2001; 2001US-0278150P.

PR 18-APR-2001; 2001US-0284591P.

PR 23-APR-2001; 2001US-0285718P.

PR 19-JUN-2001; 2001US-0293327P.

PR 16-AUG-2001; 2001US-0312902P.

XX

XX (CURA-) CURAGEN CORP.

XX

XX Li L, Padigar M, Ballinger RA, Kekuda R, Colman SD, Sciore P;

PI Smithson G, Peyman JA, Macdougall JR, Stone D, Vernet CM, Shenoy S;

PI Gunther E, Miller I, Tchernev VT, Anderson D, Gusev V, Malyankar UM;

PI Zhong H, Ellerman KE, Wolenc A;

XX

XX WPI; 2002-557660/59.

DR P-PSDB; ABP51577.

XX

XX New isolated human G-protein coupled receptor X (GPCRX) polypeptide,

PT useful for treating or preventing GPCR-associated disorders e.g.

PT diabetes, atherosclerosis, cancer or obesity.

XX

XX Claim 8; Page 88; 354pp; English.

XX

XX ABQ88354 to ABQ88417 represent human G protein coupled receptor (GPCR)

CC cDNA sequences, and ABP51560 to ABP51624 represent human GPCR proteins

CC from the present invention. GPCR sequences can have neuroprotective,

CC nootropic, anti-HIV, antiasthmatic, antiarteriosclerotic, cytosstatic,

CC immunomodulator, antiinflammatory, antidiabetic, anorectic, haemostatic,

CC antibacterial, fungicide, protozoal, virucide, nephrotropic, osteopathic,

CC cardiant, antiulcer, antiallergic, hepatotropic and antiparkinsonian

CC activities, and can be used in vaccines and gene therapy. GPCR proteins

CC nucleic acid molecules, and antibodies from the present invention can be

CC used for manufacturing a medicament for treating or preventing a GPCR-

CC associated disorder or syndrome related to cell signal processing and

CC metabolic pathway modulation, such as cardiomyopathy, atherosclerosis,

CC diabetes, cancer, obesity, infections (bacterial, fungal, protozoal or

CC viral), HIV, asthma, Parkinson's disease, osteoporosis, Crohn's disease,

CC ulcers, allergies, cirrhosis, glomerulonephritis, stroke, systemic lupus

CC erythematosus, or haematopoietic disorders. Anti-GPCR antibodies can be

CC used diagnostically to monitor protein levels in tissues as part of a

CC clinical testing procedure such as in determining the efficacy of a given

CC treatment regimen. ABQ88418 to ABQ88639 represent PCR primers and probes

CC for the human GPCRs of the present invention

XX

SQ Sequence 998 BP; 190 A; 287 C; 223 G; 298 T; 0 U; 0 Other;

Query Match 99.4%; Score 980.6; DB 6; Length 998;

Best Local Similarity 99.6%; Pred. No. 6.5e-308;

Matches 983; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 ATGTCCAGCACTCTTGGCCACACATGGAATCTCTCATCAGCTGATGTGACCCCTCT 60

Db 2 ATGTCCAGCACTCTTGGCCACACATGGAATCTCTTAATCAGCTGATGTGACCCCTCT 61

Qy 61 GTCTTCTCTCTGGGCATCCAGGCTCGAACAATTCATTTGTGGCTCTCACTCCCT 120

Db 62 GTCTTCTCTCTGGGCATCCAGGCTCGAACAATTCATTTGTGGCTCTCACTCCCT 121

Qy 121 GTGTGTGGCTTAGGCACAGCCACAAATTTGTGGCAATATAAATCTCTGTGTGTGTGCC 180

Db 122 GTGTGTGGCTTAGGCACAGCCACAAATTTGTGGCAATATAAATCTCTGTGTGTGTGCC 181

Qy 181 ACTGAACCAAGTCTTGGCAAGCCCTGTGTA CTTTTCTGTGCATGCTCTCAACCATGCAC 240

Db 182 ACTGAACCAAGTCTTGGCAAGCCCTGTGTA CTTTTCTGTGCATGCTCTCAACCATGCAC 241

Qy 241 TTGGCTGCTCTGTCTCCACAGTTCGCCAAGCTACTGGCTATCTTCTGTGTGGAGCCGGA 300

Db 242 TTGGCTGCTCTGTCTCCACAGTTCGCCAAGCTACTGGCTATCTTCTGTGTGGAGCCGGA 301

Db	122	GTGTGTGGCTTAGGCAACAGCCACAATTTGTGGGCAATAAATACTATCTGTGTGTGTGTGTC	181
Qy	181	ACTGAACCAAGTCTTGGACAAGACCTGTGTACCTTTTTCTGTGTCATGCTCTCAACCATCGAC	240
Db	182	ACTGAACCAAGTCTTGGACAAGACCTGTGTACCTTTTTCTGTGTCATGCTCTCAACCATCGAC	241
Qy	241	TTGGCTGCCTCTGTCTTCCACAGTTCCCAAGCTACTGCGCTATCTTTCTGGTGTGGAGCCGGA	300
Db	242	TTGGCTGCCTCTGTCTTCCACAGTTCCCAAGCTACTGCGCTATCTTTCTGGTGTGGAGCCGGA	301
Qy	301	CATATATCTGCCTCTGCCTGCCTGGCCACATATGTTCTTCAATTCATGCCCTTCGCAATGATG	360
Db	302	CATATATCTGCCTCTGCCTGCCTGGCCACAGATGTTCTTCAATTCATGCCCTTCGCAATGATG	361
Qy	361	GAGTCCACTGTGTACTGTGGCCATGGCCCTTTGATCGCTACGTGGCCATCTGCCACCCACTC	420
Db	362	GAGTCCACTGTGTACTGTGGCCATGGCCCTTTGATCGCTACGTGGCCATCTGCCACCCACTC	421
Qy	421	CGCTATGCCAATTCCTCACTGACACCATCATTCGCCACATAGGGGTGGCAGCTGTAGTG	480
Db	422	CGCTATGCCAATTCCTCACTGACACCATCATTCGCCACATAGGGGTGGCAGCTGTAGTG	481
Qy	481	CGAGCTCCCTGTCTCATGCTCCCATGTCCCTCTTTATTTGGGCGTTTGAACCTTTCGCAA	540
Db	482	CGAGCTCCCTGTCTCATGCTCCCATGTCCCTCTTTATTTGGGCGTTTGAACCTTTCGCAA	541
Qy	541	AGCCATGTGATCCTTACACAGCTACTGTGAGACATGCGCTGTGGTGAAGCTGGCCTGTGGA	600
Db	542	AGCCATGTGATCCTTACACAGCTACTGTGAGACATGCGCTGTGGTGAAGCTGGCCTGTGGA	601
Qy	601	GACACCAAGCGCTAACCGGTGTGTATGGGTGACAGCTGCACCTGTGGTCAATTTGGGGTGTGAC	660
Db	602	GACACCAAGCGCTAACCGGTGTGTATGGGTGACAGCTGCACCTGTGGTCAATTTGGGGTGTGAC	661
Qy	661	TTGTTTTCGATGTGPTCTCCTATAGCCCTAAATTGCACAAGCTGTCTTGGCTCTCATATCC	720
Db	662	TTGTTTTCGATGTGPTCTCCTATAGCCCTAAATTGCACAAGCTGTCTTGGCTCTCATATCC	721
Qy	721	CATGAAGTCCGTTCCAAAGCCCTTAGGGACCTGTGGTTCCCATGTCTGTGTCACTCCTCATC	780
Db	722	CATGAAGTCCGTTCCAAAGCCCTTAGGGACCTGTGGTTCCCATGTCTGTGTCACTCCTCATC	781
Qy	781	TCTTATACACCAAGCCCTCTTCTCCTTTTTTACACACCGCTTTTGGCCATCACGTTCCAGTGC	840
Db	782	TCTTATACACCAAGCCCTCTTCTCCTTTTTTACACACCGCTTTTGGCCATCACGTTCCAGTGC	841
Qy	841	CATATTTCATTTCTTTTGGCCAAATGTTTATCTGCTTTTGGCACCTGCTCTTAACTCCTG	900
Db	842	CATATTTCATTTCTTTTGGCCAAATGTTTATCTGCTTTTGGCACCTGCTCTTAACTCCTG	901
Qy	901	GTATATGGAGTTTAAAGACCAACAGATCCGTAAAGAGTTGTCAAGGTTGTTTCAAAGTGGG	960
Db	902	GTATATGGAGTTTAAAGACCAACAGATCCGTAAAGAGTTGTCAAGGTTGTTTCAAAGTGGG	961
Qy	961	CAGGGAATGGGCATCAAGGCATCTGAG	987
Db	962	CAGGGAATGGGCATCAAGGCATCTGAG	988

RESULT 7
ADC86344
ID ADC86344 standard; DNA; 1390 BP.
XX
AC ADC86344;

XX
DT 01-JAN-2004 (first entry)
XX
XX Human GPCR gene SEQ ID NO:797.
XX
XX
KW ds; gene; human; GPCR;
KW guanosine triphosphate-binding protein coupled receptor; gene therapy.
XX
XX
XX Homo sapiens.
CS

XX	PN	EP1270724-A2.
XX	XX	02-JAN-2003.
XX	PD	
XX	PF	18-JUN-2002; 2002EP-00013517.
XX	XX	
XX	PR	18-JUN-2001; 2001JP-00246789.
XX	XX	
XX	PA	(NAAD-) NAT INST ADVANCED IND SCI & TECHNOLOGY.
XX	PA	(ADSC-) CENT ADVANCED SCI & TECHNOLOGY INCUBATIO.
XX	PI	Suwa M, Asai K, Akiyama Y, Aburatani H;
XX	XX	
XX	DR	WPI; 2003-315783/31.
XX	DR	P-PSDB; ADC86345.
XX	XX	
PT	PT	New polynucleotide, useful for preparing a composition for treating a
PT	PT	patient in need of increased or suppressed activity or expression of the
PT	PT	guanosine triphosphate-binding protein coupled receptor.
XX	XX	
XX	PS	Claim 1; SEQ ID NO 797; 28pp; English.
XX	CC	The invention relates to a novel polynucleotide encoding a guanosine
XX	CC	triphosphate-binding protein coupled receptor (GPCR). A polynucleotide of
XX	CC	the invention may have a use in gene therapy. The polynucleotide and
XX	CC	polypeptide are useful for preparing a composition for treating a patient
XX	CC	in need of increased or suppressed activity or expression of the
XX	CC	guanosine triphosphate-binding protein coupled receptor. The
XX	CC	polynucleotide sequences shown in ADC85548-ADC87616 encode GPCR's of the
XX	CC	invention.
XX	XX	
XX	SQ	Sequence 1390 BP; 317 A; 353 C; 293 G; 427 T; 0 U; 0 Other;
	Query Match	99.4%; Score 980.6; DB 9; Length 1390;
	Best Local Similarity	99.6%; Pred. No. 7,9e-308;
	Matches 983; Conservative	0; Mismatches 4; Indels 0; Gaps 0;
QY	1	ATGTCACGACCTTTGGGCACAACATGGAATCTCTCATCACACTGATGTTGACCCCTTCT 60
DB	201	ATGTCACGACCTTTGGGCACAACATGGAATCTCTTAATCACACTGATGTTGACCCCTTCT 260
QY	61	GTCCTTCTTCCTCGGCATCCAGGCTCGGAACAATTTCATTGTGGCTCTCACTCCCT 120
DB	261	GTCTTCTTCCTCGGCATCCAGGCTCGGAACAATTCATTGTGGCTCTCACTCCCT 320
QY	121	GTGTGTGGCTTAGGCACAGCCACAATTGTGGCAATATAACTATATCTGTGTGTGTGTGCC 180
DB	321	GTGTGTGGCTTAGGCACAGCCACAATTGTGGGCAATATAACTATATCTGTGTGTGTGTGCC 380
QY	181	ACTGAACGAGCTTTGCACAGGCTGTACCTTTTCTGTGATGCTCTCAACATCGAC 240
DB	381	ACTGAACGAGCTTTGCACAGGCTGTGTACCTTTTCTGTGATGCTCTCAACATCGAC 440
QY	241	TTGGCTGCTCTGTCTCCACAGTCCCAAGCTACTGGCTATCTTCTGTGTGTGGAGCGGA 300
DB	441	TTGGCTGCTCTGTCTCCACAGTCCCAAGCTACTGGCTATCTTCTGTGTGTGGAGCGGA 500
QY	301	CATATATCTGCTCTGCTCGCTCGCTGGCACATATGTTCTTCATTCATGCTTCTGCATGATG 360
DB	501	CATATATCTGCTCTGCTCGCTCGCTGGCACATATGTTCTTCATTCATGCTTCTGCATGATG 560
QY	361	GAGTCCACTGTGCTACTGCGCATGGCTTTGATCGTACGTGCCATCTGCCACCCACTC 420
DB	561	GAGTCCACTGTGCTACTGCGCATGGCTTTGATCGTACGTGCCATCTGCCACCCACTC 620
QY	421	CGCTATGCCACAATCTCTCACTGACACCATCATTTGCCCAATAGGGGTGGCAGCTGTAGTG 480
DB	621	CGCTATGCCACAATCTCTCACTGACACCATCATTTGCCCAATAGGGGTGGCAGCTGTAGTG 680
QY	481	CGAGGCTCCCTGCTCATGCTCCCATGTCCTCTTTATTGGGCGTTTGAACTTCTGCGCA 540
DB	681	CGAGGCTCCCTGCTCATGCTCCCATGTCCTCTTTATTGGGCGTTTGAACTTCTGCGCA 740

865 GTTATCTGCTTTGGCCACCTGCTCTTAATCTCTGFGTATATGAGTTAAGACCAAAACAG 924
841 GTTATCTGCTTTGGCCACCTGCTCTTAATCTCTGFGTATATGAGTTAAGACCAAAACAG 900
925 ATCCGTAAGAGTTGTCAGGGTGTTCAAAGTGGGAGGGAATGGGCATCAAGGCATCT 984
901 ATCCGTAAGAGTTGTCAGGGTGTTCAAAGTGGGAGGGAATGGGCATCAAGGCATCT 960
985 GAG 987
961 GAG 963
RESULT 9
ID ABK16633
XX ABK16633 standard; cDNA; 966 BP.
AC ABK16633;
XX
DT 14-MAR-2002 (first entry)
XX
DE Human G-coupled receptor (GCRC) cDNA, Seq ID No 42.
XX
KW Human; cytotstatic; neuroprotective; immunosuppressant; nootropic;
anti-inflammatory; anti-viral; gastrointestinal; cardiovascular;
cerebroprotective; G-coupled receptor; cell proliferative disease;
lymphoma; leukaemia; breast cancer; cirrhosis; neurological disorder;
stroke; Alzheimer's disease; multiple sclerosis; mental retardation;
cardiovascular disease; atherosclerosis; angina pectoris; indigestion;
congestive heart failure; gastrointestinal disorder; dysphagia; AIDS;
gastroitis; autoimmune disorder; inflammatory disorder; Crohn's disease;
systemic lupus erythematosus; metabolic disorder; diabetes; obesity;
viral infection; herpesvirus; parvovirus;
acquired immune deficiency syndrome; ss.
XX
OS Homo sapiens.
XX
XX WO200190359-A2.
XX
XX 29-NOV-2001.
XX
XX 22-MAY-2001; 2001WO-US016833.
XX
XX 22-MAY-2000; 2000US-0206222P.
XX
XX 25-MAY-2000; 2000US-0207476P.
XX
XX 02-JUN-2000; 2000US-0208834P.
XX
XX 02-JUN-2000; 2000US-0208861P.
XX
XX 07-JUN-2000; 2000US-0209868P.
XX
XX (INCY-) INCYTE GENOMICS INC.
XX
XX Patterson C, Tribouley CM, Yao MG, Griffin JA, Thornton M, Lu Y;
PI Kallick DA, Gandhi AR, Au-Young J;
XX
XX WPI; 2002-106199/14.
XX
XX P-PSDB; AAU080511.
XX
XX New G-protein coupled receptors useful for treating or preventing cell
proliferative (e.g. leukemia), neurological (e.g. stroke), cardiovascular
or autoimmune/inflammatory disorders.
XX
XX Claim 5; Page 146; 148pp; English.
XX
XX The invention relates to a novel human G-coupled receptor (I). (I) and
its corresponding polynucleotides are useful for diagnosing, treating or
preventing cell proliferative diseases (e.g. lymphoma, leukaemia, breast
cancer or cirrhosis), neurological disorders (e.g. stroke, Alzheimer's
disease, multiple sclerosis or mental retardation), cardiovascular
diseases (e.g. atherosclerosis, angina pectoris or congestive heart
failure), gastrointestinal disorders (e.g. dysphagia, indigestion or
gastroitis), autoimmune/inflammatory disorders (e.g. AIDS, Crohn's disease
or systemic lupus erythematosus) or metabolic disorders (e.g. diabetes or

CC obesity), or viral infections (e.g. infection by herpesvirus or
CC parvovirus). ABK16633 represent novel human G-coupled receptor
CC coding sequences of the invention
XX
SQ Sequence 966 BP; 182 A; 276 C; 217 G; 291 T; 0 U; 0 Other;
Query Match 96.9%; Score 956.6; DB 6; Length 966;
Best Local Similarity 99.6%; Pred. No. 4.1e-300;
Matches 959; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 25 ATGGAATCTCTCATCACACATGATGTTGACCCCTTCTGCTCTTCTCTCTGGCATCCCA 84
DB 1 ATGGAATCTCTCATCACACATGATGTTGACCCCTTCTGCTCTTCTCTCTGGCATCCCA 60
QY 85 GGTCTGGAACAATTTCTATTTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 144
DB 61 GGTCTGGAACAATTTCTATTTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 120
QY 145 ATTGTGGCAATATAAATCTATTTCTGTTGTTGCTCTCTCTCTCTCTCTCTCTCTCTCTCT 204
DB 121 ATTGTGGCAATATAAATCTATTTCTGTTGTTGCTCTCTCTCTCTCTCTCTCTCTCTCTCT 180
QY 205 GTGTACCTTTTCTGTGATGCTCTCAACCATCGACTTGGCTGCTCTCTCTCTCTCTCTCTCT 264
DB 181 GTGTACCTTTTCTGTGATGCTCTCAACCATCGACTTGGCTGCTCTCTCTCTCTCTCTCTCT 240
QY 265 CCCAAGCTTACTGGGTATTTCTGTTGTTGAGCGGACATATATCTGCTCTCTCTCTCTCTCT 324
DB 241 CCCAAGCTTACTGGGTATTTCTGTTGTTGAGCGGACATATATCTGCTCTCTCTCTCTCTCT 300
QY 325 GCACATATGTTCTTCATTCATGCTCTCTGATGATGAGTCCACTGCTCTCTCTCTCTCTCTCT 384
DB 301 GCACATATGTTCTTCATTCATGCTCTCTGATGATGAGTCCACTGCTCTCTCTCTCTCTCT 360
QY 385 GCCTTTGATCGCTACGTGGCCATCTGCCACCCACTCCGCTATGCCACAACTCTCTCTCTCT 444
DB 361 GCCTTTGATCGCTACGTGGCCATCTGCCACCCACTCCGCTATGCCACAACTCTCTCTCTCT 420
QY 445 ACCATCATTTGCCACATAGGGGTGGCAGTGTAGTGGAGGCTCCCTGCTCTCTCTCTCTCTCT 504
DB 421 ACCATCATTTGCCACATAGGGGTGGCAGTGTAGTGGAGGCTCCCTGCTCTCTCTCTCTCTCT 480
QY 505 TGTCTCTCTTTTATTTGGGGGTTTGAACCTTCTGCCAAAGCCATGTCATCTCTACACAGTAC 564
DB 481 TGTCTCTCTCTTTTATTTGGGGGTTTGAACCTTCTGCCAAAGCCATGTCATCTCTACACAGTAC 540
QY 565 TGTGAGCACATGGCTGTGTGAAAGCTGGCCTGTGAGAGACACCCAGGCTTAACCGTGTGTAT 624
DB 541 TGTGAGCACATGGCTGTGTGAAAGCTGGCCTGTGAGAGACACCCAGGCTTAACCGTGTGTAT 600
QY 625 GGGCTGACAGCTGACATGCTGTGGTCAATGGGGTGTGATGTTTGTGATTTGCTCTCTCTCTAT 584
DB 601 GGGCTGACAGCTGACATGCTGTGGTCAATGGGGTGTGATGTTTGTGATTTGCTCTCTCTCTAT 660
QY 685 GGCCTAATTGACACAGCTGCTCTCGCTCTCATCCCATGAAGCTCGGTCCAAAGGCCCTTA 744
DB 661 GGCCTAATTGACACAGCTGCTCTCGCTCTCATCCCATGAAGCTCGGTCCAAAGGCCCTTA 720
QY 745 GGGACCTGTGGTTCCTCATGCTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCC 804
DB 721 GGGACCTGTGGTTCCTCATGCTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCC 780
QY 805 TTTTATACACACCGCTTTGGGCCATCACGTTCCAGTCCATATTCACATTTCTTTTGGCCAAAT 864
DB 781 TTTTATACACACCGCTTTGGGCCATCACGTTCCAGTCCATATTCACATTTCTTTTGGCCAAAT 840
QY 865 GTTATCTGCTTTTGGCCACCTGCTCTTAATCTCTGTTGATATGAGTTAAGACCAAAACAG 924
DB 841 GTTATCTGCTTTTGGCCACCTGCTCTTAATCTCTGTTGATATGAGTTAAGACCAAAACAG 900
QY 925 ATCCGTAAGAGTTGTCAGGGTGTTCAAAGTGGGAGGGAATGGGCATCAAGGCATCT 984
DB 901 ATCCGTAAGAGTTGTCAGGGTGTTCAAAGTGGGAGGGAATGGGCATCAAGGCATCT 960

This Page Blank (LSP10)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 27, 2004, 16:11:37 ; Search time 93 Seconds
(without alignments)
5889.640 Million cell updates/sec

Title: US-10-081-775-1_COPY_537_1523

Perfect score: 987
Sequence: 1 atgtccagcactcttgcca.....tggscatcaaggaatctgag 987

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA.*

- 1: /cgn2_6/ptodata/2/ina/5A COMB.seq.*
- 2: /cgn2_6/ptodata/2/ina/5B COMB.seq.*
- 3: /cgn2_6/ptodata/2/ina/6A COMB.seq.*
- 4: /cgn2_6/ptodata/2/ina/6B COMB.seq.*
- 5: /cgn2_6/ptodata/2/ina/PCTUS COMB.seq.*
- 6: /cgn2_6/ptodata/2/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	334	33.8	963	4	US-09-439-313-526 Sequence 526, App
2	334	33.8	963	4	US-09-636-215-526 Sequence 526, App
3	334	33.8	963	4	US-09-685-166A-526 Sequence 526, App
4	330.8	33.5	1474	1	US-08-465-980-1 Sequence 1, Appli
5	330.8	33.5	1474	2	US-09-053-303-1 Sequence 1, Appli
6	330.8	33.5	1474	4	US-09-339-115-1 Sequence 1, Appli
7	330.8	33.5	1474	5	PCT-US95-07093-1 Sequence 1, Appli
8	326.2	33.0	1539	4	US-09-668-680-13 Sequence 13, Appl
9	142.4	14.4	1854	4	US-09-016-434-1312 Sequence 1312, Ap
10	137.2	13.9	966	3	US-08-748-506-5 Sequence 5, Appli
11	133.2	13.5	966	3	US-08-748-506-7 Sequence 7, Appli
12	131.8	13.4	966	3	US-08-748-506-6 Sequence 6, Appli
13	131	13.3	1713	2	US-08-467-948A-1 Sequence 1, Appli
14	131	13.3	1713	3	US-08-467-947A-1 Sequence 1, Appli
15	126.8	12.8	966	3	US-08-748-506-8 Sequence 8, Appli
16	118.4	12.0	1990	4	US-09-016-434-1056 Sequence 1056, Ap
17	117.6	11.9	675	4	US-09-465-901-43 Sequence 43, Appl
18	117	11.9	1282	4	US-09-016-434-1413 Sequence 1413, Ap
19	116.8	11.8	3459	4	US-09-016-434-1363 Sequence 1363, Ap
20	115.4	11.7	1035	4	US-09-546-986A-1 Sequence 1, Appli
21	115.4	11.7	1035	4	US-09-524-730-1 Sequence 1, Appli
22	109	11.0	1062	4	US-09-668-680-10 Sequence 10, Appl
23	107	10.8	678	4	US-09-465-901-45 Sequence 45, Appl
24	106.4	10.8	1080	4	US-09-668-680-9 Sequence 9, Appli
25	105.8	10.7	1351	4	US-09-546-986A-5 Sequence 5, Appli
26	105.8	10.7	1351	4	US-09-524-730-5 Sequence 5, Appli
27	102.2	10.4	1438	4	US-09-016-434-1313 Sequence 1313, Ap

28	101.2	10.3	666	4	US-09-465-901-33 Sequence 33, Appl
29	99.6	10.1	542	4	US-09-016-434-627 Sequence 627, App
30	99.4	10.1	1065	4	US-03-546-986A-7 Sequence 7, Appli
31	99.4	10.1	1065	4	US-09-524-730-7 Sequence 7, Appli
32	92.2	9.3	234	4	US-09-016-434-1047 Sequence 1047, Ap
33	91.4	9.3	675	4	US-09-465-901-17 Sequence 17, Appl
34	90.6	9.2	669	4	US-09-465-901-25 Sequence 25, Appl
35	89	9.0	669	4	US-09-465-901-11 Sequence 11, Appl
36	88.8	9.0	648	4	US-09-016-434-1374 Sequence 1374, Ap
37	87.2	8.8	648	4	US-09-016-434-1371 Sequence 1371, Ap
38	85.6	8.7	900	3	US-09-085-371-5 Sequence 5, Appli
39	84.2	8.5	645	4	US-09-016-434-1372 Sequence 1372, Ap
40	82.2	8.3	648	4	US-09-016-434-1375 Sequence 1375, Ap
41	81.6	8.3	1290	2	US-08-827-291A-1 Sequence 1, Appli
42	79.2	8.0	675	4	US-09-016-434-1373 Sequence 1373, Ap
43	78.6	8.0	648	4	US-09-016-434-1376 Sequence 1376, Ap
44	78.4	7.9	669	4	US-09-465-901-13 Sequence 13, Appl
45	75.4	7.6	952	4	US-09-016-434-1115 Sequence 1115, Ap

ALIGNMENTS

RESULT 1
US-09-439-313-526
; Sequence 526, Application US/09439313
; Patent No. 6329505

; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun

; APPLICANT: Dillon, Davin C.

; APPLICANT: Mitcham, Jennifer L.

; APPLICANT: Harlocker, Susan Louise

; APPLICANT: Jiang Yuqi

; APPLICANT: Reed, Steven G.

; APPLICANT: Kalos, Michael

; APPLICANT: Fanger, Gary

; APPLICANT: Retter, Mark

; APPLICANT: Solk, John

; APPLICANT: Day, Craig

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND

; TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER

; FILE REFERENCE: 210121.427C9

; CURRENT APPLICATION NUMBER: US/09/439,313

; NUMBER OF SEQ ID NOS: 575

; SOFTWARE: FastSeq For Windows Version 3.0

; SEQ ID NO 526

; LENGTH: 963

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-439-313-526

Query Match 33.8%; Score 334; DB 4; Length 963;
Best Local Similarity 60.2%; Pred. No. 1.6e-99;

Matches 553; Conservative 0; Mismatches 365; Indels 0; Gaps 0;

QY	63	CTTCTTCCTCGGCACTCCAGGCTCGAAACAATTTGCTGCTCTACTCCCTGT	122
DB	30	CTTTGGCTTATTGGTATCCCAAGATTAGAGAAAGCCCAATTTCTGGTTGGCTTCCCCCT	89
QY	123	GTGTGGCTTAGGCACAGCCACAATTTGTGGCAATATACTATTCTGGTTGTTGTGCCAC	182
DB	90	CTTTCCATGATGATGATGGCAATGTTTGGAACTGCATCGTGCTTCATCGTAAGGAC	149
QY	183	TGAACCCAGTCTTGCAAGCCCTGTGTACTCTTTTGTGTCATGCTCTCAACCATCGACTT	242
DB	150	GGACGCGACCTGCGAGCTCCGATGTACCTTTCTCTCATGCTTGCAGCCATTGACCT	209
QY	243	GGCTGCTCTGTCTCCACAGTTCACAGTCTCCAGTACTGGTATCTTCTGTTGGAGCCGACA	302
DB	210	GGCTTATCCACATCCACCATCCATGATCTTCTGCTTCTTCTGTTGTTGATTCGCGAGA	269
QY	303	TATATCTGCTCTGCTGCTGCTGCACATATGTTCTTCATTCATGCTTCTGCTGATGGA	362

Db 270 GATTAGCTTTGAGGCTGTCTTACCAGATGTTCTTTATTCATGCCCTCTCAGCCATTGA 329
QY 363 GTCCACTGTCTACTGGCCATGGCTTTGATGCTAGCTAGTGGCCATCTGCCACCACTCCG 422
Db 330 ATCCACCACTCTGCTGGCCATGGCTTTGACCGTTATGTGGCCATCTGCCACCACTCCG 389
QY 423 CTATGCCACAATCTCTCACTGACACCATCATTCGCCACATAGGGTGGCAGCTGTAGTGG 482
Db 390 CCAATGCTGAGTGTCTCAACAATACAGTAACAGCCCAAGATTGGCATCGTGGCTGTGTCGG 449
QY 483 AGGCTCCCTGTCTCATGCTCCCATGTCCTTTCTTTATTTGGGCGTTTGAACCTTCTGCCAAG 542
Db 450 CGGATCCCTCTTTTTCCTCCACTGCTCTGCTGATCAAGCGCTGGCCCTTCTGCCACTC 509
QY 543 CATGATGATCTTACACAGTACTGTGAGCAGATGCTGTGGTGAAGCTGGCTGTGGAGA 602
Db 510 CAAATGCTCTCTCGCATCTCTATTTGTCTCCACAGGATGTAATGAAGTTGGCCCTATGCAGA 569
QY 603 CACCAAGGCTTAACCGTGTGTATGGGCTGACAGCTGCACCTGTTGGTCAATTTGGGTTGACTT 662
Db 570 CACTTTGGCCAAATGTGGTATATGTTCTTACTGCACTTCTGCTGGTCAATGGGCTGGAGCT 629
QY 663 GTTTTGCATGTCTCTCTATGCTCCCTAATTCACAAAGCTGTCTTGGCTCTCATPCCA 722
Db 630 AATGTTCACTCTCTTGTCTCTATTTCTGTATAATACGAACGGTCTGCAACTGCCCTTCCA 689
QY 723 TGAAGCTGGTCAAGGCTTGAAGCACTGTGGTTCGCCATGCTGTGTCATCCTCATCTC 782
Db 690 GTCAAGCGGGCAAGGCTTTGGAACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 749
QY 783 TTATACACAGCCCTCTCTCTCTTTTACACACCGCTTTGGCCATCAGTTCAGTCCA 842
Db 750 CTATGTGCCACTTATTTGGCTCTCAGTTGTATACCGCTTTGGAAACAGCTTCATPCCAT 809
QY 843 TATTACAACTCTTTTGGCCAAATGTTATCTGCTTTTGGCCACTGCTCTTAATPCTGTGT 902
Db 810 TGTGCTGT 869
QY 903 ATATGGAGTTAAGCAACAGATCCGTAAGAGTTGTCAAGGTGTGTCAAGGTGGGCA 962
Db 870 CTATGTGCCAAACCAACAGATCAGAACCGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 929
QY 963 GGAATGGGCATCAAGGC 980
Db 930 TGACAAGGACTTGCAGGC 947

RESULT 2

US-09-636-215-526

; Sequence 526, Application US/09636215

; Patent No. 6620922

; GENERAL INFORMATION:

; APPLICANT: Xu, Jiangchun

; APPLICANT: Dillon, Devin C.

; APPLICANT: Mitcham, Jennifer L.

; APPLICANT: Harlocker, Susan L.

; APPLICANT: Jiang, Yuqi

; APPLICANT: Henderson, Robert A.

; APPLICANT: Kalos, Michael D.

; APPLICANT: Fanger, Gary R.

; APPLICANT: Retter, Marc W.

; APPLICANT: Stolk, John A.

; APPLICANT: Day, Craig H.

; APPLICANT: Vedvick, Thomas S.

; APPLICANT: Carter, Derrick

; APPLICANT: Li, Samuel

; APPLICANT: Wang, Aijun

; APPLICANT: Skeiky, Yasir A.W.

; APPLICANT: Hepler, William

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND

; TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER

; FILE REFERENCE: 210121.42717C17

; CURRENT APPLICATION NUMBER: US/09/636,215

; CURRENT FILING DATE: 2000-08-10

; NUMBER OF SEQ ID NOS: 852

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 526

; LENGTH: 963

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-636-215-526

Query Match 33.8%; Score 334; DB 4; Length 963;
Best Local Similarity 60.2%; Pred. No. 1.6e-99;
Matches 553; Conservative 0; Mismatches 365; Indels 0; Gaps 0;

QY 63 CTTCTCTCTCTGGGCATCCCAAGGTCTGGAAACAATTTATTTGTGGCTCTCACTCTCTCT 122
Db 30 CTTTGTGCTTATTTGATTTCCCAAGGATTAGAGAAAGCCCAATTTCTGGTGGCTTCCCT 89
QY 123 GTGTGGCTTAGGCACAGCCACAATTTGTGGGCAATATACTATTCTGGTGTGTGGCCAC 182
Db 90 CTTTCCATGTATGTAGTGGCAATGTTTGGAAACTGCATCGGTCTTTCATCGTAAGGAC 149
QY 183 TGAACCAAGTCTTGCCACAAGCCTGTGTACCTTTTCTGTGCATGCTCTCAACCATGACT 242
Db 150 GGAACGCAAGCTGCAAGCTCCGATGTACTCTTCTGTGCATGCTTGGCCATTTGATTC 209
QY 243 GGCTGCTCTGTCTCCACAGTTCCTCAAGTCTATGGCTATCTTCTGTGTGGAGCCGACA 302
Db 210 GGCTTATCCCAATCCACATCCATGCTTAAGATCTTCCCTTTTCTGTGTGGTATTC 269
QY 303 TATATCTGCTCTGCTGCTGGCCTGGACATATGTTCTTCAATGAGCTTCTGATGATGA 362
Db 270 GATTAGCTTTGAGGCTCTGTCTTACCCAGATGTTCTTTATTCATGCCCTCTCAGCCA 329
QY 363 GTCCACTGTGTACTTGGCCATGGCTTTGATGCTGCTAGCTGGCCATCTGCCACCACTCG 422
Db 330 ATCCACCACTCTGTGGCCATGGCTTTGACGTTATGTGGCCATCTGCCACCACTGG 389
QY 423 CTATGCCACAATCTCTCACTGACACCATCATTTGCCACATAGGGGTGGCAGCTGTAGTGG 482
Db 390 CCATGCTGAGTGTCAACAATAACAGTAAAGCCAGATGGCATCGTGGCTGTGTGCTCG 449
QY 483 AGGCTCCCTGCTCATGCTCCCATGCTCCCTTCTTTATTTGGGGCTTTGAACCTTCTG 542
Db 450 CGGATCCCTCTTTTTCCTCCACTGCTCTGCTGATCAAGCGGCTGGGCTTCTTGCCTC 509
QY 543 CCATGTGATCTTACACAGCTACTGTGACACATGGCTGTGTGGAAGCTGGCTGTGGAGA 602
Db 510 CAATGTCTCTCGCACTCTTATTGTGCCACAGGATGTAATGAAGTTGGCTGTATGCAGA 569
QY 603 CACCAGGCTTAACCGTGTGTATGGGCTGACAGCTGCACTGTGTGTGCTATTTGGGTTG 662
Db 570 CACTTTGCCCAATGTGGTATATGCTTACTGTCCATTTCTGTGTGCTATGGGCTGGAGCT 629
QY 663 GTTTTGCATTTGCTCTCTCTATGCTCCCTTAATTTGCAAGCTGCTCTTCCCTCTCAT 722
Db 630 AATGTTCACTCTCTGTGCTTATTTTCTGATAATACGAACCGTTCTGCAACTGCTTCCA 689
QY 723 TGAAGCTCGTCCCAAGGCTCTAGGCACTGTGGTTCCTCATGCTGTGTGCTCATCTC 782
Db 690 GTCAAGCGGGCCAAAGGCTTTGGAACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 749
QY 783 TTATACACAGCCCTCTTCTCTTTTACACAGCTTTGGCCATCAGGCTTCAGTCCA 842
Db 750 CTATGTGCCACTTATTTGGCCCTCTCAGTTGTATACCGCTTTGGAAACAGCTTCAT 809
QY 843 TATTACAACTCTTTTGGCCAAATGTTATCTGCTTTTGGCCACTGCTCTTAATPCTGTGT 902
Db 810 TGTGCTGT 869
QY 903 ATATGGAGTTAAGCAACAGATCCGTAAGAGTTGTCAAGGTGTGTCAAGGTGGGCA 962
Db 870 CTATGTGCCAAACCAACAGATCAGAACCGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 929

SEQUENCE CHARACTERISTICS:

LENGTH: 1474 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cdna
FEATURE:

NAME/KEY: CDS
LOCATION: 274..1233
US-08-465-980-1

Query Match 33.5%; Score 330.8; DB 1; Length 1474;

Best Local Similarity 60.0%; Pred. No. 2.4e-98;

Matches 551; Conservative 0; Mismatches 367; Indels 0; Gaps 0;

QY 63 CTTCTTCCTCCCTGGGATCCAGGCTCGAACAATTTCAATTTGTGGCTCTCACTCCCTGT 122
Db 303 CTGTGTGCTTATGGTATCCAGGATTAGAGAACCCCAATTTCTGGGTTGGCTTCCCT 362
QY 123 GTGTGGCTTAGGCACAGCCCAATTTGTGGCAATATAACTATTCTGGTTGTTGGCCAC 182
Db 363 CTTTCCATGTATGTAGTGGCAATGTGTGAACTGCATCGTGCTTTCATCGTAAGGAC 422
QY 183 TGAACCAAGTCTTGCAACAGCTGTGTACCTTTTCTGTGCATGCTCTCAACCATGACTT 242
Db 423 GGAACGCGACCTGCGACGCTCCGATGTACCTTTCTCTGCATGCTTGACGCCATTGACCT 482
QY 243 GCGTCCCTCTGTCTCCACAGTTCACAGCTACTGGCTATCTTCTGGTGTGGAGCCGGACA 302
Db 483 GGCCTTATCCACATCCACATGCCCTTAAGATCTTGGCCCTTTCTGGTTGATTCCCGAGA 542
QY 303 TATATCTGCTCTGCTGCTGGCACATATGTTCTTTCATTCATGCTTCTGCATGATGA 362
Db 543 GATTAGCATTTAGGCGCTGTCTTACCAGATGTTCTTTATTCATGCGCTCTGAGCCATTGA 602
QY 363 GTCCAGTGTCTACTGGCCATGGCCTTTGATCGCTACGTGGCCATCTGCCACCCACTCG 422
Db 603 ATCCACCATCTGCTGGCCATGGCCTTTGACCGTTATGTGGCCATCTGCCACCCACTGG 662
QY 423 CTATGCCACAATCCTCACTGACACCATATTCGCCACATAGGGTGGGAGCTGTAGTGG 482
Db 663 CCATGCTGCTCTCGCACTCTTATGTGTCCACAGGATGTAATGAAGTTGGCCTATGCAGA 722
QY 483 AGGCTCCCTGCTCATGCTCCCATGCTCCCTTTCTTTATGGGCTTTTGAACCTTCTGCCAAG 542
Db 723 CGGATCCCTCTTTTTCCTCACTGCTCTGTGATCAAGCGCTGGCCTTCTGCCACTC 782
QY 543 CCATGTGATCTTACACAGCTGTGTGAGACATGCTGTGTGAAGCTGGCCTGTGGAGA 602
Db 783 CAATGTCTCTCGCACTCTTATGTGTCCACAGGATGTAATGAAGTTGGCCTATGCAGA 842
QY 603 CACGAGGCTTAACCGTGTGTATGGCTCACAGCTGCACCTGTGTGTCATTTGGGTTGACTT 662
Db 843 CACTTTGCCAATGTGTATATGTTCTTACTTGCCATTTCTGTGTGTCATGGGCGTGGAGCT 902
QY 663 GTTTTGCATTTGGTCTCTCTATGCCCCTAATTTGCAAGCTGTGCTTCCCTCTCATCCCA 722
Db 903 AATGTTCACTCTCTGCTCTATTTTCTGATTAATAGAACGTTCTGCAACTGCCCTTCAA 962
QY 723 TGAAGCTCGTTCGAGGCCCTTAGGACCTGTGGTTCCCATGCTGTGTGTCATCTCATCTC 782
Db 963 GTCAGAGCGGCCCAAGGCTTTTGAACCTGTGTGTGCACATGTTGTGTGTACTCGCCTT 1022
QY 783 TTATACACAGCCCTCTTCTCTCTTTTACACAGCTTTGGCCATCAGCTTCCAGTCCA 842
Db 1023 CTATGTGCCATTATGCGCTCTCAGTTGTACACCGCTTTTGAACAGCCCTTCATCCCAT 1082
QY 843 TATTACATTTTGGCCAAATTTTATCTGCTTTTGGCACCTGCTCTTAATCTCTGGT 902
Db 1083 TGTGCGTGTGTGTATGGGTGACATCTACCTGCTGCTGCTCTGCTGCTGCTGCTGCTGCT 1142
QY 903 ATATGGAAGTTAAGACCAACAGATCCGTTAAAGAGTTGTCAAGGTGTTTCAAAGTGGCA 962

Db 1143 CTATGTGTCARAAACCAACAGATCAGACGGGTGCTGGCTATGTTCAAGATCAGCTG 1202
QY 963 GGAATGGCATCAAGGC 980
Db 1203 TGACAGGACTTGACGC 1220

RESULT 5

US-09-053-303-1

; Sequence 1, Application US/09053303

; Patent No. 5948890

; GENERAL INFORMATION:

; APPLICANT: Soppet, Daniel R.

; APPLICANT: Li, Yi

; APPLICANT: Rosen, Craig A.

; APPLICANT: Ruben, Steven M.

; TITLE OF INVENTION: HUMAN G-PROTEIN RECEPTOR HPAJ70

; NUMBER OF SEQUENCES: 8

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN, CECCHI,

; ADDRESSEE: STUART & OLSTEIN

; STREET: 6 Becker Farm Road

; CITY: Roseland

; STATE: New Jersey

; COUNTRY: USA

; ZIP: 07068

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent In Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/053.303

; FILING DATE:

; CLASSIFICATION:

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/465,980

; FILING DATE: 06-JUN-1995

; ATTORNEY/AGENT INFORMATION:

; NAME: Ferraro, Gregory D.

; REGISTRATION NUMBER: 36,134

; REFERENCE/DOCKET NUMBER: 325800-446

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 201-994-1700

; TELEFAX: 201-994-1744

; INFORMATION FOR SEQ ID NO: 1:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 1474 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: cdna

; FEATURE:

; NAME/KEY: CDS

; LOCATION: 274..1233

; US-09-053-303-1

Query Match

Best Local Similarity 33.5%; Score 330.8; DB 2; Length 1474;

Matches 551; Conservative 0; Mismatches 367; Indels 0; Gaps 0;

QY 63 CTTCTTCCTCCCTGGGATCCAGGCTCGAACAATTTCAATTTGTGGCTCTCACTCCCTGT 122
Db 303 CTGTGTGCTTATGGTATCCAGGATTAGAGAACCCCAATTTCTGGGTTGGCTTCCCT 362
QY 123 GTGTGGCTTAGGCACAGCCCAATTTGTGGCAATATAACTATTCTGGTTGTTGGCCAC 182
Db 363 CTTTCCATGTATGTAGTGGCAATGTGTGAACTGCATCGTGCTTTCATCGTAAGGAC 422
QY 183 TGAACCAAGTCTTGCAACAGCTGTGTACCTTTTCTGTGCATGCTCTCAACCATGACTT 242
Db 423 GGAACGCGCTGACAGCTCCGATGTACCTTTCTCTGTCATGCTTGACGCCATTGACCT 482

Qy	243	GGCTGCCTCTGTCTCCACAGTTCCTCCAAAGCTACTGGGCTATCTTCTGGTGTGGAGCCGGACA	302
Db	483	GGCCTTATCACATCCACCATGCTAAGATCTTGGCCCTTTCTGGTTTGATTTCCGAGA	542
Qy	303	TATATCTGCCTCTGCCTGCTGGCACATATGTTCTTCAATCATGSCCTCTGCAATGATGGA	362
Db	543	GATTAGCAATTGAGGCCCTGTCTTACCAGATGTTCTTTATTCATGCCCTCTCAGCCATTGA	602
Qy	363	GTTCCACTGTGCTACTGTGCCCATGGCCTTTTGATCGCTAGTGGGCCATCTGCCACCACCTCCG	422
Db	603	ATCCACCATCCTGCTGGCCATGGCCTTGACCGTTATGTGGCCATCTGCGCACCCACTGCG	662
Qy	423	CTATGCCACAATCCTCACTGACACATCATTTGCCCCACATAGGGGTGGCAGCTGTAGTGGG	482
Db	663	CCATGCTGCAGTGTCTCAACAATAACAGTAAACAGCCAGATTGGCATCTGGCTGTGGTCCG	722
Qy	483	AGGCTCCCTGCTCATGTCTCCCATGTGCCCTCTTTATTGGGGGTTTGAATCTTCTGCCAAAG	542
Db	723	CGGATCCCTCTTTTTTTTCCCACTGSCCTCTGCTGATCAAGCGGCTGGCCCTTCTGCCACTC	782
Qy	543	CCATGTGATCCTACACACGTA CTGTGAGCACATGGCTGTGTGAAGCTGGCCTGTGGAGA	602
Db	783	CAATGTCTCTCGCACCTCTCTATTGTGTGCCACAGGATGTAATGAAGTTGGCCTATGCAGA	842
Qy	603	CACAGGCGCTAACCGTGTGTATGGGCTGACAGCTGCACTGTTGTGTCTAATGGGGTTGACTT	662
Db	843	CACTTTGGCCCAATGTGGTATATGGTCTTACTGCCATTTCTGCTGGCTATGCGCGTGGACGT	902
Qy	663	GTTTTGCAATTGGTCTCTCCATATGCCCTAAATTGCAACAGCTGTCCCTCGCCTCTCATCCCCA	722
Db	903	AATGTTTCACTCTCTGTGCTCTATTTTCTGATTAATACGAACGGTCTTGGCACTGCCCTCCAA	962
Qy	723	TGAAGCTCGGTCCAGGCCCTTAGGGACCTGTGGTTTCCCATCTCTGTGTCTCATCTCATCTC	782
Db	963	GTCAAGGGGCGCAAGGCCCTTGGAACTGTGTGTCAACATTGGTGTGGTACTCGCCTT	1022
Qy	783	TTATACACAGCCCTCTTCTCCTTTTWTATACACCGCTTTTGGCCCATCACTGTTCCAGTCCA	842
Db	1023	CTATGTGCCACTTATTGGCCCTCTCAGTTGTACACCGCTTTTGGAAACAGCCCTTCATCCCAT	1082
Qy	843	TATTACATTTCTTTTGGCCAAATGTTATCTGCTTTTGGCACCTGCTCTTAATCTGTGGT	902
Db	1083	TGTGCGTGTCTGATGGGTGCATCTACTCTGCTCTGCTCTCTGTCATCAATCCCATCAT	1142
Qy	903	ATATGGAGTTAAGACCAACCAAGATCCGTAAGAAGAGTTGTCAAGGTGTTTCAAAGTGGGCA	962
Db	1143	CTATGGTGCCTAAACCAACAGATCAGAACAGGGGTGTGCTATGTTCAAGATCAGGTG	1202
Qy	963	GGGAATGGGCATCAAGC	980
Db	1203	TGACAAGGACTTGCAGGC	1220

RESULT 6

```

US-09-339-115-1
; Sequence 1, Application US/09393115
; Patent No. 6372891
; GENERAL INFORMATION:
; APPLICANT: Soppet, Daniel R.
; APPLICANT: Li, Yi
; APPLICANT: Rosen, Craig A.
; APPLICANT: Ruben, Steven M.
; TITLE OF INVENTION: HUMAN G-PROTEIN RECEPTOR HPRAJ70
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN, CECCHI
; ADDRESSEE: STUART & OLSTEIN
; STREET: 6 Becker Farm Road
; CITY: Roseland
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07068
; COMPUTER READABLE FORM:

```

QY 663 GTTTTCATGGTCTCTCTATGCCCTAAATGACAAAGCTGTCTTGGCTCTCATCCCA 722
Db |||||
QY 903 AATGTTCACTCTCTTCTCTATTTCTGATTAATACGAACGGTTCTGCAACTGCTTCCAA 962
Db |||||
QY 723 TGAAGTCGGTCAAGGCCCTAGGACCTGTGTTTCCCATGTCTGTGTGTCATCTCATCTC 782
Db |||||
QY 963 GTCAGAGCGGGCAAGGCTTTTGAACCTGTGTGTACACATTTGGTGTGTTACTCGCTT 1022
Db |||||
QY 783 TTATACACAGCCCTCTCTCTCTTTTACACACCGCTTTGGCCATCAGCTTCCAGTCCA 842
Db |||||
QY 1023 CTATGTGCCACTTATTTGGCTCTCTAGTTGTACACCGCTTTTGAACAGAGCTTCTATCCCAT 1082
QY 843 TATTACACATCTTTTGGCCAAATGTTTATCTGCTTTTGGCCACCTGCTCTTAATCCTGTGGT 902
Db |||||
QY 1083 TGTGCGTGTGTGTCATGGGTGACATCTACTGCTGTGCTGCTCTCTGTCATCAATCCCATCAT 1142
QY 903 ATATGGAGTTAAGACCAACAGATCCGTAAAGAGTTGTACGGGTGTTTCAAAGTGGGCA 962
Db |||||
QY 1143 CTATGGTGCACAAACCAACAGATCAGAACACGGGTGCTGGCTATGTTCAAGATCAGCTG 1202
QY 963 GGAATGGGCATCAAGGC 980
Db |||||
QY 1203 TGACAGGACTTGCAGGC 1220
Db |||||

RESULT 7

PCT-US95-07093-1
; Sequence 1, Application PC/TUS9507093
; GENERAL INFORMATION:
; APPLICANT: Soppet, Daniel R.
; APPLICANT: Li, Yi
; APPLICANT: Rosen, Craig A.
; APPLICANT: Ruben, Steven M.
; TITLE OF INVENTION: HUMAN G-PROTEIN RECEPTOR HPRAJ70
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN, CECCHI,
; ADDRESS: STUART & OLSTEIN
; STREET: 6 Becker Farm Road
; CITY: Roseland
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07068
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/07093
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/465,980
; FILING DATE: 06-JUN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Ferraro, Gregory D.
; REGISTRATION NUMBER: 36,134
; REFERENCE/DOCKET NUMBER: 325800-446
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-994-1700
; TELEFAX: 201-994-1744
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1474 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 274..1233

PCT-US95-07093-1

Query Match 33.5%; Score 330.8; DB 5; Length 1474;
Best Local Similarity 60.0%; Pred. No. 2.4e-98;
Matches 551; Conservative 0; Mismatches 367; Indels 0; Gaps 0;
QY 63 CTTCTCTCTCTCTGGGCATCCAGGCTCTGGAACAATTTCAATTTGTGGCTCTCACTCCCTGT 122
Db CTGTGTGCTTATTTGTTATCCAGGATTAGAAGAACCCATTTCTGGGTTGGCTTCCCTCT 362
QY 123 GTGTGGCTTAGGCACAGCCACAAATTTGTGGCAATATACTATTCTGTGTTGTGGTCCAC 182
Db CTTTTCATGATGATGTAGTGGCAATGTGTGGAACCTGTCATCTGTTCTTCAATGTAAGGAC 422
QY 183 TGAACAGCTTTTGCAACAGCTGTGTACCTTTTCTGTGTCATGCTCTCAACCATGACATT 242
Db GGAACGACGCTGTCAGCTTCCAGTGTACCTTTCTGTGTCATGCTGTGACCATTTGACCT 482
QY 243 GGCTGCTCTGTCTCCACAGTTCCCAAGCTACTGGCTATCTTCTGGTGTGGAGCCGGACA 302
Db GSCCTTATCCACATCCACATGCCCTTAAGATCTTCCCTTTTCTGGTTGATTCCCGAGA 542
QY 303 TATATCTGCTCTGCTGCGCACATATGTTCTTCAATTCATGCTCTTCGATGATGGA 362
Db GATTAGCATTTGAGGCTCTTACCCAGATGTTCTTTATTATGAGCTCTCAGGCCATTGA 602
QY 363 GTCCACTGTGTACTGCGCATGGCCTTTGATGCTGCTACGTGGCCATCTGCCACCCACTCG 422
Db ATCCACCATCTCTGTGGCCATGGCCTTTGACCGTTATGTGGCCATCTGCCACCCACTCG 662
QY 423 CTATGCCACAATCCTCACTGACACATCATTTGCCACATAGGGGTGGCAGCTGTAGTGG 482
Db CCAATGCTCGAGTGTCTCAACAATACAGTAACAGCCACAGATTTGGCATCTGTGGTGTGG 722
QY 483 AGGCTCCCTGTCTCATGCTCCCATGTCCTTTCTTTATTTGGGGTGTGAACCTTCTGCAAA 542
Db CGGATCCCTCTTTTTCCTCTCTGCTGATCAAGGGCTGGGCTTCTGCACTC 782
QY 543 CCATGTGATCTTACACAGCTACTGTGAGCAATGCTGTGTGAGCTGGCTGTGGAGA 602
Db CAATGTCTCTCGCACTCTTATTTGTCTCCACAGGATGTAATGAAGTTGGCTATGCAGA 842
QY 603 CACGAGGCTAACCGTGTGTATGGCTGACAGCTGCACTGTTGCTCATTTGGGTTGACTT 862
Db CACTTTGCCCAATGTGTATATGGTCTTACTGCAATCTCTGTGTTGATGGGCGTGGACGT 902
QY 663 GTTTTGCATTTGCTCTCTCTATGCCCCTAAATGACAAAGCTGTCTTGGCTCTCATCCCA 722
Db AATGTTCACTCTCTTGTCTTATTTCTGATAATAGGAACGGTTCTGCAACTGCTTCCAA 962
QY 723 TGAAGCTCGGTCCAAAGGCCCTAGGACCTGTGGTTCCCATGCTGTGTGTCATCTCTCATCTC 782
Db GTCAGAGCGGGCCAAAGGCTTTTGAACCTGTGTGTACACATTTGGTGTGTGTTACTGCGCTT 1022
QY 783 TTATACACAGCCCTCTCTCTCTTTTACACACCGCTTTTGGCCATCAGTTCCAGTCCA 842
Db CTATGTGCCACTTATTTGGCTCTCAGTTGTACACCGCTTTTGGAAACAGGCTTCTATCCCAT 1082
QY 843 TATTACATCTTTTGGCCAAATGTTTATCTGCTTTTGGCCACCTGCTCTTAATCTCTGGT 902
Db TGTGCGTGTGTGTCATGGGTGACATCTACCTGCTGCTGCTGCTCTCTGTCATCAATCCATCAT 1142
QY 903 ATATGGAGTTAAGACCAACAGATCCCGTAAAGAGTTGTGAGGGTGTTCAGAGTGGGCA 962
Db CTATGGTGCACAAACCAACAGATCAGAACACGGGTGCTGGCTATGTTCAAGATCAGCTG 1202
QY 963 GGAATGGGCATCAAGGC 980
Db |||||
QY 1203 TGACAGGACTTGCAGGC 1220
Db |||||

RESULT 8

US-09-668-680-13

Db	1177	ATTGTCATGGGCTAACTGTGGCTCTGTCTGGCCATGGGACTGGATTCCATTCTCAATTGCC	1233
QY	676	CTCTCTCTATGGCCCTAATTGCACAAGCTGTCTCTTGGCTCTCATCCCATGAAGCTGGTCC	735
Db	1237	ATTTCCTATGGCTTTATCTCTCATGCAAGTCTTTTACGCTTCCATCTCATGATGCCAGCAC	1296
QY	736	AAGGCCCTAGGGACCTGTGGTTCCCATGTGTGTGCATCTCTCATCTCTTATACACAGGCC	795
Db	1297	AAAGCTCTGAGTACCTGTGGCTCCCATGTGGCATCATCTCTGTTTCTACATCCCTGCC	1356
QY	796	CTCTCTCTCTTTTATACACACCGCTTTGGCCATCAC---GTTCCAGTCCATATTACATT	852
Db	1357	TTCCTCTCTCTCTCAACCACCGCTTTGGTCACCAAGTCCCAAGCATGTGCATC	1416
QY	853	CTTTTGGCAAGTATTATCTGCTTTTGGCACTGCTCTTAATCTCTGGTATATGGATT	912
Db	1417	TTTCTGGCTAAATCTCTATGTGTGGTGGCTCTCTGTATCAATCTATTCTCTATGGAGCT	1476
QY	913	AAGACCAACAGATCCGTAACAGAGTTGTCCAGGTGTTTCCAAAGTGGGCAG	963
Db	1477	AGAACCAAGAGATCGGAGTCGACTCTTAAACTGCTTCCCTCCCTGGGGAAG	1527
RESULT 9			
US-09-016-434-1312			
; Sequence 1312, Application US/09016434			
; Patent No. 6500938			
; GENERAL INFORMATION:			
; APPLICANT: Janice Au-Young			
; APPLICANT: Jeffrey J. Seilhamer			
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING			
; TITLE OF INVENTION: PATHWAY GENE EXPRESSION			
; NUMBER OF SEQUENCES: 1490			
; CORRESPONDENCE ADDRESS:			
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.			
; STREET: 3174 PORTER DRIVE			
; CITY: PALO ALTO			
; STATE: CALIFORNIA			
; COUNTRY: USA			
; ZIP: 94304			
; COMPUTER READABLE FORM:			
; MEDIUM TYPE: Floppy disk			
; COMPUTER: IBM PC compatible			
; OPERATING SYSTEM: PC-DOS/MS-DOS			
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2			
; CURRENT APPLICATION DATA:			
; APPLICATION NUMBER: US/09/016,434			
; FILING DATE: HERewith			
; CLASSIFICATION:			
; PRIOR APPLICATION DATA:			
; APPLICATION NUMBER:			
; FILING DATE:			
; CLASSIFICATION:			
; ATTORNEY/AGENT INFORMATION:			
; NAME: Zeller, Karen J.			
; REGISTRATION NUMBER: 37,071			
; REFERENCE/DOCKET NUMBER: PA-0002 US			
; TELECOMMUNICATION INFORMATION:			
; TELEPHONE: (650) 855-0555			
; TELEFAX: (650) 845-4166			
; INFORMATION FOR SEQ ID NO: 1312:			
; SEQUENCE CHARACTERISTICS:			
; LENGTH: 1854 base pairs			
; TYPE: nucleic acid			
; STRANDEDNESS: single			
; TOPOLOGY: linear			
; IMMEDIATE SOURCE:			
; LIBRARY: GENBANK			
; CLONE: g32085			
US-09-016-434-1312			
Query Match		14.4%;	Score 142.4; DB 4; Length 1854;

	Best Local Similarity	49.6%;	Pred. No. 2.6e-36;	
	Matches	365;	Conservative	0; Mismatches 371; Indels 0; Gaps 0;
QY	58	TCGTGCTTCTTCTCCTCTGGGCATCCAGGCTCGGAACAATTTCAITTTGGCTCTCACTC	117	
Db	347	TCAGACTTCTGCTCTCTGGGCTCGCCCATCCAACAGAGCAGAAACCTGTGCTATGCC	406	
QY	118	CCTGTGTGGCTTAGGCACAGCCACAATTTGTGGCAATAAATATTTCGTGTGTTGTT	177	
Db	407	CTGTTCTTTGGCATGTATCTTACCACTCTCTGGGAACCTCCTCATCTATGTCTCAT	466	
QY	178	GCCACTGAACAGTCTTGGACAAGCCTGTGTACCTTTTCTGTGCACTCTCTCAACCATC	237	
Db	467	CGACTGGACTCCCATCTCCACAGCCTATGTATTTCTCAGCAACTTGTCTCTCTCT	526	
QY	238	GACTTGGCTGCTCTGCTCCACAGTTCGCAAGCTCTGTAACCTTTTCTGTGCACTCTCT	297	
Db	527	GACCTCTGCTCTCTTCCTGACCAATCCCAAGTTGTTACAGAACTGACAGAACCGAC	586	
QY	298	GGACATATATCTGCTCTGCTGCTGTCGCACATATGTTCTTCAITCAATGCTTCTGCATG	357	
Db	587	CCATCCATCCCTATGCGGACTGCTGACCCCAATGTACTTCTCTGTTATTTGGAGAC	646	
QY	358	ATGGAGTCCAATGTCTACTGGCCATGGCCCTTTGATCGTACGTGGCCATCTGCCACCCA	417	
Db	647	CTGGAGAGCTTCTCCTTGTGGCCATGGCCCTATGACCGCTATGGCCATCTGCTTCCCC	706	
QY	418	CTCCGCTATGCCAACAATCTCTACTGACACCATCATTTGCCACATAGGGGTGGCAGCTGA	477	
Db	707	CTGCACCTACACGGCCATCATAGACCCCATGCTCTGTCTGCGCCTGGTGGCGCTGCTCG	766	
QY	478	GTGCGAGGCTCCCTGCTCATGCTCCCATGTCCTTCTTATTGGGCGTTTGAATCTTGC	537	
Db	767	GTGCTGACCACCTTCCATGCCAATGTTTACACATTTACTCATGGCCAGGTTGTGTTTTGT	826	
QY	538	CAAAAGCCATGTGATCCTTACACACATCTGTGAGCACATGGCTGTGTGGAAGCTGGCCCTGT	597	
Db	827	GCAGACAATGTGATCCGCCACTTTTCTGTGATATGTCTGCTGTGTGAAGCTGGCCCTTC	886	
QY	598	GGAGACACCGGCTTAACCGTGTGTATGGCTGACAGTGCATGTTGGTCAITGGGGGTT	657	
Db	887	TCTGACACTCGAGTTAATGAATGGGTGATATTTATCATGGAGGGCTCATCTTTGTCA	946	
QY	658	GACTTGTGTTGTCATGTGCTCTCTATGGCCCTAATTGCAAGCTGTCTTTCGCTCTCA	717	
Db	947	CCATTCCTACTCATCTCTGGGTCTATGCAAGAAATGCTCTCTCCATCTCTCAAGTCCCT	1006	
QY	718	TCCCATGAAGCTCGGTCCTCAAGGCCCTAGGACACCTGTGGTTCCTCCATGCTCTGTC	777	
Db	1007	TCTTCTAAGGGTATCTGCAAGGCCTTCTTACTTTGGGCTCCCACTGTCTGTGGTGTCA	1066	
QY	778	ATCTCTTATACACAG	793	
Db	1067	CTGTTCTATGGAACCG	1082	

```

RESULT 10
US-08-748-506-5
; Sequence 5, Application US/08748506
; Patent No. 6159707
; GENERAL INFORMATION:
; APPLICANT: Ronnett et al.
; TITLE OF INVENTION: NOVEL SPERM RECEPTORS
; NUMBER OF SEQUENCES: 31
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Leydig, Voit & Mayer, Ltd.
; STREET: Two Prudential Plaza, Suite 4900
; CITY: Chicago
; STATE: IL
; COUNTRY: US
; ZIP: 60601-6780
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
;

```

```

;
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/748,506
; FILING DATE: 08-NOV-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/033,751
; FILING DATE: 09-NOV-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; REFERENCE/DOCKET NUMBER: 74940
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-616-5600
; TELEFAX: 312-616-5700
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 966 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; US-08-748-506-5

      Query Match      13.9%; Score 137.2; DB 3; Length 966;
      Best Local Similarity 49.2%; Pred. No. 8.9e-35;
      Matches 389; Conservative 0; Mismatches 395; Indels 6; Gaps 1;

QY 138 AGCCACAATTGGGGCAATATACTATTCGTGTTGTTGGTCCACTGAACCAAGTCTTGCA 197
Db 136 AGTATCACTAACAGAAATACTCTCTAGTCCCTTGCTATTGTACCAAGTCCATCTCTACA 185

QY 198 CAAGCCTGTGACCTTTTCTCTGTGCATGCTCTCAACCATCGACTGGCTGGCTCTGTC 257
Db 186 CACCCCAATGTACTTCTTTTCGGCCAACTGTCTCTCTGGAGATTGGCTATCTTCTGTC 245

QY 258 CACAGTTCCCAAGCTACTGGGTATCTCTGTGTGGAGCGCGAATATATCTTGCCCTCGC 317
Db 246 TGTCAATACCCAAAGATGCTGCAGAGCCTTGTGAGTGAGCGCCAGAGAGATCTCTCGGAGGG 305

QY 318 CTGCCTGSCACATATGTTTCTTTCATTGCTCTCTGCATGATGAGTCCACTGTGCTACT 377
Db 306 ATGTGCCACACAGATGTTTTTTTTTCGCAATTTTGTGTATACTAGTGTGCCATTTGGC 365

QY 378 GGCCATGGCCTTTGATCGCTACGTGGCCATCTGCACCCACTCCGGCTATGCCAATCCCT 437
Db 366 AGCCATGGCCTTTGACCGCTGATGCCATATGCTCCCACTCCCACTATGCAACCCGAAT 425

QY 438 CACTGCACACCATATTCGCCACATAGGGGTGGCAGCTGTAGTGGAGGCTCCCTGTCTCAT 497
Db 426 GAGTCGNAGAGGTATGTGCCCAATTGGCAATTTGTTTCATGGGGAATGGGATGCTAGTAAG 485

QY 498 GCTCCCATGTCCCTTCTTTATTATGGGCGTTTGAACCTCTGCGCAAGCCATGTGATCCTACA 557
Db 486 TCTGGGCAAAACCAATTTTATTTTCTCTTGAACCTCTGTGGACCCCTGTGAATAGACCA 545

QY 558 CACGTACTGTGAGCACATGGCTGTGFGAAGCTGGCGCTGTGGAGACACCAAGGCTTAACCG 617
Db 546 CTTCTCTGTGACCTTCACCTCTCTCTGGCATTGCTGTGGAGATACATCCCAAAACGA 605

QY 618 TGTGATGGGCTGACAGCTGCACTGTGTGTCATTGGGGTTGACTTGTTTTGCATTTGGTCT 677
Db 606 GGCTGCCACTTTGTGTGTAGCAGTCTCTCTGCAATCTAGCCCAATTTTCTGTCATCTTA 665

QY 678 CTCCTATGCCCTTAATTGSCAACAAGCTGTCTCTTCGCCCTCTCATCCCATGAAGCTCGGTCCA 737
Db 666 TTCTATGTCAAATTTCTCATTTTCAGTGTCTNCTGATGCCCTTCACTGAGGGGCCATAA 725

QY 738 GGCCTTAGGACCTGTGGTTCATGTCGTGTGTCATCTCATCTCTTATACACAGCCCT 797
Db 726 AGCTCTTTCACCTGTTCGTCTCACCTACTTGTAGTCACACTTTTATGGCTTCAGCATG 785

```


QY 599 GAGACACAGCCCTAACGGTGTGTATGGGCTGACAGCTGCACTGTGTGGTCAATGGGGTGG 658
Db 693 CTGATACCTGGCTCAACAGGTGTGTATCTTTGAAGCCTGATGTTTCATCTGTGGGAC 752
QY 659 ACTTGTGTTGCAATGGTCTCTCTATGCGCCCTAAATGACAAAGCTGTCTTGGCCTCTCAT 718
Db 753 CACTCTGCTGGTGTCTCTCTACTACTACACATCTCTGGGGGACATCTGAGGATCCAGT 812
QY 719 CCCATGAAGCTCGGTCCAAAGCCCTAGGGACCTGTGGTTCCTCATGTCTGTGTCAATCTCA 778
Db 813 CTGGGAGGGCCGAGAAAGCCCTTCTCCACCTGTCTCTCCACCTCTGCGTAGTGGGAC 872
QY 779 TCTCTTATACACAGCCCT 829
Db 873 TCTCTTTTGSAGCGCCATCGTATGATGATGCGCCCTTAAGTCCCGGCATC 923

RESULT 14

US-08-467-947A-1
; Sequence 1, Application US/08467947A
; Patent No. 6090575
; GENERAL INFORMATION:
; APPLICANT: LI, YI
; APPLICANT: CAO, LIANG
; APPLICANT: NI, JIAN
; APPLICANT: GENTZ, REINER
; APPLICANT: BULT, CAROL J.
; APPLICANT: SUTTON III, GRANGER G.
; APPLICANT: ROSEN, CRAIG A.
; TITLE OF INVENTION: Polynucleotides Encoding Human G-Protein
; TITLE OF INVENTION: Coupled Receptor GPR1
; NUMBER OF SEQUENCES: 30
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.
; STREET: 1100 NEW YORK AVE., NW, SUITE 600
; CITY: WASHINGTON
; STATE: DC
; COUNTRY: USA
; ZIP: 20005

COMPUTER READABLE FORM:
; MEDIUM TYPE: FLOPPY DISK
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PATENTIN RELEASE #1.0, VERSION #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/467,947A
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 435

PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/04079
; FILING DATE: 30-MAR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: STEFFE, ERIC K.
; REGISTRATION NUMBER: 36,688
; REFERENCE/DOCKET NUMBER: 1488.1140002/EKS/KLM
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-371-2600
; TELEFAX: 202-371-2540

INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1713 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: both
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 116..1003
US-08-467-947A-1

Query Match 13.3%; Score 131; DB 3; Length 1713;
Best Local Similarity 48.1%; Pred. No. 1.4e-32;
Matches 371; Conservative 0; Mismatches 400; Indels 0; Gaps 0;

QY 59 CTGCTCTTCTCTCTCTGGGCAATCCAGGCTGGGAACAATTTCAATTTTGGTCTCTCACTCC 118
Db 153 CAGAGTTCTCTCTCTCTCTCTGGGATTTCTCTGGGCCCCAAGGATTCAGATGCTCTCTTGGGC 212
QY 119 CTGTGTCTGGCTTAGGCACAGCCACAATTTGGGCAATATAAATATTTCTGTGTGTGTGTG 178
Db 213 TCTTCT 272
QY 179 CCACTGAACAGCTTTGACAAAGCCTGTGACCTTTTCTGTGTGATGCTCTCTCAACACATCG 238
Db 273 CACTGGACTCCAGACTCCACACACCCCATGTACTTCTCTCTCTCTCTCTCTCTCTCTCTCT 332
QY 239 ACTTGGCTGCTCTGTCTCTCCAGCTTCCAGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 298
Db 333 ACATCGGCTATGCTCTGCAACACAGTGGCCAGATGTGGTGAACCTCTCTCTCTCTCTCTCT 392
QY 299 GACATATATCTGCTCTGCTCTGCTGCGCAGACATATGTTCTTCTCATTCATGCTCTCTGCA 358
Db 393 AGCCATCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 452
QY 359 TGGAGTCCACTGTGCTACTGCGCCATGCGCTTTGATCGCTACGTGGCCATCTGCGCACCCAC 418
Db 453 CTGAATGCTCTCTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 512
QY 419 TCGCTATGCCACAATCTCTACTGACACCATCATTTGCCACATAGGCCACATAGGGGTGGCAGT 478
Db 513 TCCGATATTTTATCATCATCATGACCTTGGAAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTG 572
QY 479 TGGCAGGCTCCCTGCTCTCATGCTCCCATGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 538
Db 573 CATGTGCT 632
QY 539 AAAGCCATGTGATCTCTACACAGCTACTGTGTGAGCAGATGGCTGTGTGTGAGCTGGCCTGT 598
Db 633 GGCCTCGTGAATCAACACCTTCTCTGTGAAATCTCTGCTGCTGCTGCTGCTGCTGCTGCTG 692
QY 599 GAGACACAGCCCTAACGGTGTGTATGGGCTGACAGCTGCACTGTGTGTGTGTGTGTGTGT 658
Db 693 CTGATACCTGGCTCAACAGGTGTGTATCTTTTGAAGCCTGATGTTTCATCTCTGTGGGAC 752
QY 659 ACTTGTCTTGTGCT 718
Db 753 CACTCTGCTGGTGTGTGTCT 812
QY 719 CCCATGAAGCTCGGTCCAAAGCCCTAGGGACCTGTGGTTCCTCATGTCTGTGTCTCAATCTCA 778
Db 813 CTGGGAGGGCCGAGAAAGCCCTTCTCCACCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 872
QY 779 TCTCTTATACACAGCCCT 829
Db 873 TCTTCTTTTGSAGCGCCATCGTATGATGATGCGCCCTTAAGTCCCGGCATC 923

RESULT 15

US-08-748-506-8
; Sequence 8, Application US/08748506
; Patent No. 6159707
; GENERAL INFORMATION:
; APPLICANT: Ronnett et al.
; TITLE OF INVENTION: NOVEL SPERM RECEPTORS
; NUMBER OF SEQUENCES: 31
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Leydig, Voit & Mayer, Ltd.
; STREET: Two Prudential Plaza, Suite 4900
; CITY: Chicago
; STATE: IL
; COUNTRY: US
; ZIP: 60601-6780
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS

Job time : 96 secs

SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/748,506
FILING DATE: 08-NOV-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/033,751
FILING DATE: 09-NOV-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
REFERENCE/DOCKET NUMBER: 74940
TELEPHONE: 312-616-5600
TELEFAX: 312-616-5700
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 966 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-748-506-8

Query Match 12.8%; Score 126.8; DB 3; Length 966;
Best Local Similarity 49.5%; Pred. No. 2.3e-31;
Matches 326; Conservative 0; Mismatches 332; Indels 0; Gaps 0;

QY	170	TTGTTGTTGGCCACTGAACACGCTCTGGACAGCCCTGTGTACCTTTTCTGTGCATGCTCT	229
Db	158	TGCTATTGTACCACTCTCTACACACCCCATGTACTTCTTTCTGGCCAACTTGT	217
QY	230	CAACCATCGACTTGGCTGCCCTCTGTCTCCACAGTTCCCAAGCTACTGCTATCTTCTGGT	289
Db	218	CTCTCTGGAGATTGGCTATATCTTCTCTGTATACCCAAAGATGCTGCAGAGCTTGTGA	277
QY	290	GTGAGCGCGGACATATATCTGCCCTCTGCCCTGGGCACATATGTTCTTCAATTCATGCCCT	349
Db	278	GTGAGGCCCGGAGATCTTTCAGGTGGGATGTGCCACACAGATGTTTTTCTCATATTCT	337
QY	350	TCTGCATGATGGAGTCCACTGTCTGTACTGGCCATGGCCTTTGATCGCTACGTGGCCATCT	409
Db	338	TTGTTATTAACGTAGTGTGCCCTATTGGCAGCCATGGCCCTTTGACCGCTATATGGCTATAT	397
QY	410	GCCACCCACTCCGCTATGCCAAATCTCTCACTGACACACCATCAITGGCCACATAGGGTGG	469
Db	398	GTTCCCACTCCACTATGCAACCCGAATGAGTCTGTAGGTATGTGCCACTTGGCAATTG	457
QY	470	CAGCTGTAGTCCGAGGCTCCCTGTCTATGCTCCCATGTCCTTCTTTATGGGGCTTTGA	529
Db	458	TTTCATGGGTGATGGGATGCATAGTAGTCTGGGACACAGCAATTTTATTTCTCCTTGA	517
QY	530	ACTTCTGCCAAGCATGTGATCTCTACACAGTACTGTGACACATGGCTGTGGTGAAGC	589
Db	518	ACTTCTGTGGACCTGTGAGATAGACCACTTCTTCTGTGATCTTCCACCTCTCCTGGCAC	577
QY	590	TGGCCTGTGAGACACACAGGCTTAACCGTGTGTATGGGCTGACAGCTGCACTGTTGGTCA	649
Db	578	TTGCCTGTGGTGATACATCCAAATTGAGGCTGCCATCTTTGTGTAGTTGTCTCTGCA	637
QY	650	TTGGGGTTGACTGTTTGTGATTTGGTCTCTCTATAGCCCTTAATTGCAAGCTGTCCCTTC	709
Db	638	TATCTAGCCCTTTTCTGCTGATCATTTTATTCTTATGTAGAAATCTCGTTGCAGTGCTGG	697
QY	710	GCCTCTCATCCCAAGCTCGGTCCCAAGGCCCTTAGGGACCTGTGGTTCCTCATGCTGTG	769
Db	698	TGATGCCCTTACCTGAGGGGGCCACAAAGCCCTTCAACCTGTCTCTCCCACTACTTG	757
QY	770	TCATCTCATCTCTTATACACAGCCCTCTTCTCTCTTTTACACACCGCTTTGGCCA	827
Db	758	TAGTCACACTCTTTTATGGCTCAGGATCTGTACCTATTGAGGCCCTAAGTCTAGCCA	815

Search completed: August 27, 2004, 18:30:17

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 27, 2004, 16:19:27 ; Search time 540 Seconds
(without alignments)
8995.273 Million cell updates/sec

Title: US-10-081-775-1_COPY_537_1523
Perfect score: 987
Sequence: 1 atgtccagcactctggcca.....tgggcatcaagcattctgag 987

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 3237270 seqs, 2460713050 residues

Total number of hits satisfying chosen parameters: 6474540
Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA:*

- 1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/2/pubpna/US05_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
- 10: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/2/pubpna/US09D_PUBCOMB.seq:*
- 12: /cgn2_6/ptodata/2/pubpna/US09E_NEW_PUB.seq:*
- 13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
- 14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
- 15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:*
- 16: /cgn2_6/ptodata/2/pubpna/US10D_PUBCOMB.seq:*
- 17: /cgn2_6/ptodata/2/pubpna/US10E_NEW_PUB.seq:*
- 18: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
- 19: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	987	100.0	1567	15	US-10-081-775-1
2	980.6	99.4	990	9	US-09-886-055-262
3	980.6	99.4	990	10	US-09-804-291-262
4	980.6	99.4	990	13	US-10-343-650A-353
5	980.6	99.4	998	15	US-10-025-806-35
6	980.6	99.4	998	15	US-10-025-806-37
7	980.6	99.4	1366	15	US-10-017-161-923
8	980.6	99.4	1390	16	US-10-292-798-797
9	956.6	96.9	966	17	US-10-297-021-42
10	376.2	38.1	980	11	US-09-844-861A-25
11	376.2	38.1	1345	15	US-10-017-161-879
12	376.2	38.1	1986	9	US-09-864-761-3771
13	375.2	38.0	945	9	US-09-886-055-90
14	375.2	38.0	945	10	US-09-804-291-90

15	375.2	38.0	945	13	US-10-182-822A-25	Sequence 25, Appl
16	375.2	38.0	945	13	US-10-343-650A-161	Sequence 161, App
17	375.2	38.0	945	14	US-10-032-106-5	Sequence 5, Appli
18	374.6	38.0	1345	16	US-10-292-798-757	Sequence 757, App
19	373.6	37.9	945	16	US-10-387-629-161	Sequence 161, App
20	369.8	37.5	945	9	US-09-864-761-20537	Sequence 20537, A
21	368.4	37.3	1013	15	US-10-025-806-25	Sequence 25, Appl
22	368.4	37.3	1360	15	US-10-017-161-947	Sequence 947, App
23	368.4	37.3	1360	16	US-10-292-798-821	Sequence 821, App
24	367.2	37.2	982	11	US-09-844-861A-21	Sequence 21, Appl
25	366.6	37.1	1345	15	US-10-017-161-877	Sequence 877, App
26	365	37.0	1345	16	US-10-292-798-755	Sequence 755, App
27	364	36.9	1151	11	US-09-844-861A-19	Sequence 19, Appl
28	364	36.9	1151	11	US-09-886-055-86	Sequence 86, Appl
29	363.4	36.8	1400	9	US-09-804-291-86	Sequence 86, Appl
30	362.4	36.7	945	13	US-10-343-650A-159	Sequence 159, App
31	356.6	36.1	960	9	US-09-886-055-96	Sequence 96, Appl
32	356.6	36.1	960	10	US-09-804-291-96	Sequence 96, Appl
33	356.6	36.1	960	10	US-10-343-650A-317	Sequence 317, App
34	356.6	36.1	960	16	US-10-387-629-159	Sequence 159, App
35	355.6	36.0	990	13	US-10-343-650A-315	Sequence 315, App
36	355.6	36.0	990	16	US-10-387-629-157	Sequence 157, App
37	355.6	36.0	1114	17	US-10-467-252-93	Sequence 93, Appl
38	355.6	36.0	1390	15	US-10-017-161-945	Sequence 945, App
39	355.6	36.0	1390	16	US-10-292-798-819	Sequence 819, App
40	355.6	35.6	1076	13	US-10-182-822A-23	Sequence 23, Appl
41	351.2	35.3	1695	15	US-10-145-586-34	Sequence 34, Appl
42	348.2	35.3	948	15	US-10-145-586-36	Sequence 36, Appl
43	348	35.3	1300	16	US-10-292-798-833	Sequence 833, App
44	345.2	35.0	948	16	US-10-387-629-189	Sequence 189, App
45	340.4	34.5				

ALIGNMENTS

RESULT 1

US-10-081-775-1
; Sequence 1, Application US/10081775
; Publication No. US2003060409A1
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: POLYNUCLEOTIDE ENCODING A NOVEL HUMAN G-PROTEIN COUPLED
; TITLE OF INVENTION: RECEPTOR, HGPBMY25, EXPRESSED HIGHLY IN IMMUNE-RELATED
; FILE REFERENCE: D0126 NP
; CURRENT APPLICATION NUMBER: US/10/081,775
; PRIOR FILING DATE: 2002-02-21
; PRIOR APPLICATION NUMBER: US 60/270,134
; PRIOR FILING DATE: 2001-02-21
; PRIOR APPLICATION NUMBER: US 60/278,952
; PRIOR FILING DATE: 2001-03-27
; NUMBER OF SEQ ID NOS: 69
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 1567
; TYPE: DNA
; ORGANISM: homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (537)..(1523)
US-10-081-775-1

Query Match	100.0%	Score 987;	DB 15;	Length 1567;
Best Local Similarity	100.0%	Pred. No. 2.6e-307;		
Matches 987;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	ATGTCCAGCACTCTTGGCCCAACATCGAATCTCTCATCATCATGATGTGACCTTCT	60	
Db	537	ATGTCCAGCACTCTTGGCCCAACATCGAATCTCTCATCATCATGATGTGACCTTCT	596	
Qy	61	GTCTTCTTCTCTCTGGGCGATCCAGGCTCGAACAAATTCATTTGTGCTCTCTACTTCT	120	
Db	597	GTCTTCTTCTCTCTGGGCGATCCAGGCTCGAACAAATTCATTTGTGCTCTCTACTTCT	656	

```
121 GTGTGGCTTAGGCACAGCCACAAATGTGGGCAATATAAATAATCTGTGGTGTGGTGGCC 180
122 GTGTGGCTTAGGCACAGCCACAAATGTGGGCAATATAAATAATCTGTGGTGTGGTGGCC 716
181 ACTGAACAGTCTTGGCAGAGCCGTGTACCTTTCTGTGCATGCTCTCAACCATCGAC 240
182 GTGTGGCTTAGGCACAGCCGTGTACCTTTCTGTGCATGCTCTCAACCATCGAC 776
183 ACTGAACAGTCTTGGCAGAGCCGTGTACCTTTCTGTGCATGCTCTCAACCATCGAC 776
241 TTGGCTGGCTCTGTCTCCACAGTCTCCCAAGTACTGGCTATCTTCTGTGTGGAGCCGGA 300
777 TTGGCTGGCTCTGTCTCCACAGTCTCCCAAGTACTGGCTATCTTCTGTGTGGAGCCGGA 836
301 CATATATCTGGCTCTGTCTCCACAGTCTCCCAAGTACTGGCTATCTTCTGTGTGGAGCCGGA 360
837 CATATATCTGGCTCTGTCTCCACAGTCTCCCAAGTACTGGCTATCTTCTGTGTGGAGCCGGA 896
361 GAGTCCAGTGTCTTGGCCATGAGCCCTTGTATCGCTAGCTGGCCATCTGCCACCATCTC 420
897 GAGTCCAGTGTCTTGGCCATGAGCCCTTGTATCGCTAGCTGGCCATCTGCCACCATCTC 956
421 CGCTATGCCACAATCTCTCACTGACACCATCAATTTGCCACATAGGGGTGGAGCTGTAGTG 480
957 CGCTATGCCACAATCTCTCACTGACACCATCAATTTGCCACATAGGGGTGGAGCTGTAGTG 1016
481 CGAGGCTCCCTGCTCATGCTCCCATGCTCCCTTTCTTTATTTGGGCGTTTGAACCTTCCG 540
1017 CGAGGCTCCCTGCTCATGCTCCCATGCTCCCTTTCTTTATTTGGGCGTTTGAACCTTCCG 1076
541 AGCCATGTGATCTTACACAGTGTGTGAGCAGTGTGTGCTGCTGCTGCTGCTGCTGCTG 600
1077 AGCCATGTGATCTTACACAGTGTGTGAGCAGTGTGTGCTGCTGCTGCTGCTGCTGCTG 1136
601 GACACAGCCCTAACCGTGTGTATGGGCTGACAGCTGTGCTGCTGCTGCTGCTGCTGCTG 660
1137 GACACAGCCCTAACCGTGTGTATGGGCTGACAGCTGTGCTGCTGCTGCTGCTGCTGCTG 1196
661 TTGTTTGTGATGTCTTCTTATGCTTATGCTTATGCTTATGCTTATGCTTATGCTTATG 720
1197 TTGTTTGTGATGTCTTCTTATGCTTATGCTTATGCTTATGCTTATGCTTATGCTTATG 1256
721 CATGAAGCTCGTGTCCAGCCCTTAGGAGCTGTGTGCTGCTGCTGCTGCTGCTGCTGCTG 780
1257 CATGAAGCTCGTGTCCAGCCCTTAGGAGCTGTGTGCTGCTGCTGCTGCTGCTGCTGCTG 1316
781 TCTTATACACAGCCCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 840
1317 TCTTATACACAGCCCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1376
841 CATATTCATCTTTTGGCCATGTTTATCTGCTTTTGGCCATGCTTCTTATATCTGCTG 900
1377 CATATTCATCTTTTGGCCATGTTTATCTGCTTTTGGCCATGCTTCTTATATCTGCTG 1436
901 GTATATGGAGTTAAGACCAACAGATCCGTAAAGAGTTGTCTGAGGTGTTTCAAGTGGG 960
1437 GTATATGGAGTTAAGACCAACAGATCCGTAAAGAGTTGTCTGAGGTGTTTCAAGTGGG 1496
961 CAGGGAATGGGCATCAAGGCATCTGAG 987
1497 CAGGGAATGGGCATCAAGGCATCTGAG 1523
```

```
RESULT 2
US-09-886-055-262
; Sequence 262, Application US/09886055
; Patent No. US20020132273A1
; GENERAL INFORMATION:
; APPLICANT: STREYER, LUBERT
; APPLICANT: ZOZULYA, SERGEY
; TITLE OF INVENTION: RECEPTOR FINGERPRINTING, SENSORY PERCEPTION, AND
; TITLE OF INVENTION: BIOSENSORS OF CHEMICAL SENSANTS
; FILE REFERENCE: 078003-0277150
; CURRENT APPLICATION NUMBER: US/09/886,055
```

```
; CURRENT FILING DATE: 2001-06-22
; PRIOR APPLICATION NUMBER: 60/211,812
; PRIOR FILING DATE: 2000-06-22
; NUMBER OF SEQ ID NOS: 522
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 262
; LENGTH: 990
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-886-055-262

Query Match      99.4%; Score 980.6; DB 9; Length 990;
Best Local Similarity 99.6%; Pred. No. 2.3e-305;
Matches 983; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 ATGTCCAGCACTCTTGGCCACAAATGGAATCTCTCTATCACACTGATGTGGACCTTCT 60
DB 1 ATGTCCAGCACTCTTGGCCACAAATGGAATCTCTCTATCACACTGATGTGGACCTTCT 60
QY 61 GTCTTCTTCTCTGGGCATCCAGGCTGTGAAACAAATTTTCTGTTGGCTCTCACTCCCT 120
DB 61 GTCTTCTTCTCTGGGCATCCAGGCTGTGAAACAAATTTTCTGTTGGCTCTCACTCCCT 120
QY 121 GTGTGGCTTAGGCACAGCCACAAATGTGGCAATAATACTATTTCTGTTGGTGGCC 180
DB 121 GTGTGGCTTAGGCACAGCCACAAATGTGGCAATAATACTATTTCTGTTGGTGGCC 180
QY 181 ACTGAACAGTGTGTGCAAGCCCTGTGACCTTTTCTGTGATGCTCTCAACCATCGAC 240
DB 181 ACTGAACAGTGTGTGCAAGCCCTGTGACCTTTTCTGTGATGCTCTCAACCATCGAC 240
QY 241 TTGGCTGTGCTCTGTCTCCACAGTCTCCCAAGTACTGGCTATCTTCTGTTGGAGCCGGA 300
DB 241 TTGGCTGTGCTCTGTCTCCACAGTCTCCCAAGTACTGGCTATCTTCTGTTGGAGCCGGA 300
QY 301 CATATATCTGCTCTGCTGTGCTGTGCAATATGTTCTTCTTCTGTTGGTGGTGGTGG 360
DB 301 CATATATCTGCTCTGCTGTGCTGTGCAATATGTTCTTCTTCTGTTGGTGGTGGTGG 360
QY 361 GAGTCCAGTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 420
DB 361 GAGTCCAGTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 420
QY 421 CGCTATGCCACAATCTCTCACTGACACCATCAATTTGCCACATAGGGGTGGAGCTGTG 480
DB 421 CGCTATGCCACAATCTCTCACTGACACCATCAATTTGCCACATAGGGGTGGAGCTGTG 480
QY 481 CGAGGCTCCCTGCTCATGCTCCCATGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 540
DB 481 CGAGGCTCCCTGCTCATGCTCCCATGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 540
QY 541 AGCCATGTGATCTTACACAGTGTGTGAGCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTG 600
DB 541 AGCCATGTGATCTTACACAGTGTGTGAGCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTG 600
QY 601 GACACAGCCCTAACCGTGTGTATGGGCTGACAGCTGTGCTGTTGTTGCTATGTTGGGTTG 660
DB 601 GACACAGCCCTAACCGTGTGTATGGGCTGACAGCTGTGCTGTTGTTGCTATGTTGGGTTG 660
QY 661 TTGTTTGTGATGTTGCTCTCTATGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 720
DB 661 TTGTTTGTGATGTTGCTCTCTATGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 720
QY 721 CATGAAGCTCGGTCCCAAGGCCCTTAGGAGCTGTGTTGCCATGCTGTGCTGTGCTGTGCT 780
DB 721 CATGAAGCTCGGTCCCAAGGCCCTTAGGAGCTGTGTTGCCATGCTGTGCTGTGCTGTGCT 780
QY 781 TCTTATACACAGCCCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 840
DB 781 TCTTATACACAGCCCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 840
QY 841 CATATTCATCTTTTGGCCATGTTTATCTGCTTTTGGCCATGCTGCTTCTTCTTCTTCTG 900
DB 841 CATATTCATCTTTTGGCCATGTTTATCTGCTTTTGGCCATGCTGCTTCTTCTTCTTCTG 900
```



```
Db      841  CATATTACATCTTTTGGCCAAATGTTTATCTGCTTTTGGCCACTGCTCTTAATCTGTG 900
QY      901  GTATATGGAGTTAAGACCAACAGATCCGTAAAGAGTTGTCAAGGTGTTTCAAAGTGG 960
Db      901  GTATATGGAGTTAAGACCAACAGATCCGTAAAGAGTTGTCAAGGTGTTTCAAAGTGG 960
QY      961  CAGGGAATGGCATCAAGGCATCTGAG 987
Db      961  CAGGGAATGGCATCAAGGCATCTGAG 987

RESULT 3
US-09-804-291-262
; Sequence 262, Application US/09804291
; Publication No. US20030088059A1
; GENERAL INFORMATION:
; APPLICANT: ZOZULA, SERGEY
; TITLE OF INVENTION: HUMAN OLFACTORY RECEPTORS AND GENES ENCODING SAME
; FILE REFERENCE: P 0278005
; CURRENT APPLICATION NUMBER: US/09/804,291
; PRIOR FILING DATE: 2001-03-13
; PRIOR APPLICATION NUMBER: 60/188,914
; PRIOR FILING DATE: 2000-03-13
; PRIOR APPLICATION NUMBER: 60/192,033
; PRIOR FILING DATE: 2000-03-24
; PRIOR APPLICATION NUMBER: 60/198,474
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/199,335
; PRIOR FILING DATE: 2000-04-24
; PRIOR APPLICATION NUMBER: 60/207,702
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/213,849
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: 60/226,534
; PRIOR FILING DATE: 2000-08-16
; PRIOR APPLICATION NUMBER: 60/230,732
; PRIOR FILING DATE: 2000-09-07
; PRIOR APPLICATION NUMBER: 60/266,862
; PRIOR FILING DATE: 2001-02-07
; NUMBER OF SEQ ID NOS: 529
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 262
; LENGTH: 990
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-804-291-262

Query Match      99.4%; Score 980.6; DB 10; Length 990;
Best Local Similarity 99.6%; Pred. No. 2.3e-305;
Matches 983; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY      1  ATGTCCAGCACTCTTGGCCCAACAGTGAATCTCTCATCACTGATGTTGACCCCTTCT 60
Db      1  ATGTCCAGCACTCTTGGCCCAACAGTGAATCTCTCATCACTGATGTTGACCCCTTCT 60
QY      61  GTCTTCTTCTCTCTGGGCATCCAGTCTCGAACAATTTCAATTTGTTGGCTCTCACTCCCT 120
Db      61  GTCTTCTTCTCTCTGGGCATCCAGTCTCGAACAATTTCAATTTGTTGGCTCTCACTCCCT 120
QY      121  GTGTGTGGCTTATGAGCACAGCCCAATTTGGGCAATATACTATTTCTGGTTGTTTGGCC 180
Db      121  GTGTGTGGCTTATGAGCACAGCCCAATTTGGGCAATATACTATTTCTGGTTGTTTGGCC 180
QY      181  ACTGAACAGCTTTCACACAGCTGTGTACCTTTTCTGTGCATGCTCTCAACCATCGAC 240
Db      181  ACTGAACAGCTTTCACACAGCTGTGTACCTTTTCTGTGCATGCTCTCAACCATCGAC 240
QY      241  TTGGCTGCTCTGTCTCCACAGTTCCCAAGTACTGGCTATCTTCTGTGTGGAGCCGGA 300
Db      241  TTGGCTGCTCTGTCTCCACAGTTCCCAAGTACTGGCTATCTTCTGTGTGGAGCCGGA 300
QY      301  CATATATCTGCTCTGCTGCTGGCACATATGTTCTTCATTCATGCTCTCTGCATGATG 360
|||||
```

```
Db      301  CATATATCTGCTCTGCTGCTGGCACAGATGTTCTTCTATTCATGCTCTTCTGCATGATG 360
QY      361  GAGTCCACTGTGCTACTTGGCCCATGGCTTTTGATCGCTACGTGCGCCATCTGSCCACTC 420
Db      361  GAGTCCACTGTGCTACTTGGCCCATGGCTTTTGATCGCTACGTGCGCCATCTGSCCACTC 420
QY      421  CGCTATGCCACAATCTCTCACTGACACCATCATTTGCCCATAGGGGTGGCAGCTGTAGTG 480
Db      421  CGCTATGCCACAATCTCTCACTGACACCATCATTTGCCCATAGGGGTGGCAGCTGTAGTG 480
QY      481  CGAGGCTCCCTGCTCATGCTCCCATGCTCCCTTCTTTTATTTGGGCGTTTGAATCTTCCCAA 540
Db      481  CGAGGCTCCCTGCTCATGCTCCCATGCTCCCTTCTTTTATTTGGGCGTTTGAATCTTCCCAA 540
QY      541  AGCCATGTGATCCTTACACACGTACTGTGAGCACATGGCTGTGTGAAGCTGGCTCTGTGA 600
Db      541  AGCCATGTGATCCTTACACACGTACTGTGAGCACATGGCTGTGTGAAGCTGGCTCTGTGA 600
QY      601  GACACAGGCTTAACCGTGTGTATGGGTGACAGCTGCACTGTTGGTCAATGGGTTGAC 660
Db      601  GACACAGGCTTAACCGTGTGTATGGGTGACAGCTGCACTGTTGGTCAATGGGTTGAC 660
QY      661  TTGTTTTCATTTGCTCTCTCTATGCCCTAAATTGACAGCTGTCTTGGCTCTCATCC 720
Db      661  TTGTTTTCATTTGCTCTCTCTATGCCCTAAATTGACAGCTGTCTTGGCTCTCATCC 720
QY      721  CATGAAGCTCGGTCCAAGGCCCTTAGGGACCTGTGGTTCATATGCTGTGTCTCATCTCATC 780
Db      721  CATGAAGCTCGGTCCAAGGCCCTTAGGGACCTGTGGTTCATATGCTGTGTCTCATCTCATC 780
QY      781  TCTTATACACAGCCCTCTCTCTCTTTTACACACCGCTTTGGCCATCACTCCAGTTC 840
Db      781  TCTTATACACAGCCCTCTCTCTCTTTTACACACCGCTTTGGCCATCACTCCAGTTC 840
QY      841  CATATTACATTTCTTTTGGCCAAATGTTTATCTGCTTTTGGCCACTGCTCTTAATCTGTG 900
Db      841  CATATTACATTTCTTTTGGCCAAATGTTTATCTGCTTTTGGCCACTGCTCTTAATCTGTG 900
QY      901  GTATATGGAGTTAAGACCAACAGATCCGTAAAGAGTTGTCAAGGTGTTTCAAAGTGGG 960
Db      901  GTATATGGAGTTAAGACCAACAGATCCGTAAAGAGTTGTCAAGGTGTTTCAAAGTGGG 960
QY      961  CAGGGAATGGGCATCAAGGCATCTGAG 987
Db      961  CAGGGAATGGGCATCAAGGCATCTGAG 987
```

```
RESULT 4
US-10-343-650A-353
; Sequence 353, Application US/10343650A
; Publication No. US20040067499A1
; GENERAL INFORMATION:
; APPLICANT: HAGA, TATSUYA
; TITLE OF INVENTION: NOVEL G-PROTEIN COUPLED RECEPTOR
; FILE REFERENCE: 31671-186347
; CURRENT APPLICATION NUMBER: US/10/343,650A
; CURRENT FILING DATE: 2003-07-21
; PRIOR APPLICATION NUMBER: JP 2000/237818
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: JP 2001/34434
; PRIOR FILING DATE: 2001-02-13
; NUMBER OF SEQ ID NOS: 694
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 353
; LENGTH: 990
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(990)
US-10-343-650A-353
```

Query Match 99.4%; Score 980.6; DB 13; Length 990;

Best Local Similarity 99.6%; Pred. No. 2.3e-305;		US-10-025-806-35	
Matches 983; Conservative 0; Mismatches 4; Indels 0; Gaps 0;		; Sequence 35; Application US/10025806	
		; Publication No. US2003019855A1	
		; GENERAL INFORMATION:	
		; APPLICANT: Li, Li	
		; APPLICANT: Padigaru, Muralidhara	
		; APPLICANT: Ballinger, Robert	
		; APPLICANT: Kekuda, Ramesh	
		; APPLICANT: Colman, Steven	
		; APPLICANT: Spytek, Kimberly	
		; APPLICANT: Casman, Stacie	
		; APPLICANT: Edinger, Shlomit	
		; APPLICANT: Gerlach, Valerie	
		; APPLICANT: Sciore, Paul	
		; APPLICANT: Smithson, Glenda	
		; APPLICANT: Feyman, John	
		; APPLICANT: MacDougall, John	
		; APPLICANT: Stone, David	
		; APPLICANT: Vernet, Corine	
		; APPLICANT: Shenoy, Suresh	
		; APPLICANT: Gunther, Erik	
		; APPLICANT: Millet, Isabelle	
		; APPLICANT: Tchernev, Velizar	
		; APPLICANT: Anderson, David	
		; APPLICANT: Gusev, Vladimir	
		; APPLICANT: Malyankar, Uriel	
		; APPLICANT: Zhong, Haihong	
		; APPLICANT: Ellerman, Karen	
		; APPLICANT: Wolenc, Adam	
		; TITLE OF INVENTION: NOVEL	
		; FILE REFERENCE: 21402-224 AB	
		; CURRENT APPLICATION NUMBER: US/10/025,806	
		; CURRENT FILING DATE: 2001-12-19	
		; PRIOR APPLICATION NUMBER: 60/256,635	
		; PRIOR FILING DATE: 2000-12-18	
		; PRIOR APPLICATION NUMBER: 60/259,743	
		; PRIOR FILING DATE: 2001-01-04	
		; PRIOR APPLICATION NUMBER: 60/299,327	
		; PRIOR FILING DATE: 2001-06-19	
		; PRIOR APPLICATION NUMBER: 60/261,498	
		; PRIOR FILING DATE: 2001-01-12	
		; PRIOR APPLICATION NUMBER: 60/263,689	
		; PRIOR FILING DATE: 2001-01-24	
		; PRIOR APPLICATION NUMBER: 60/276,464	
		; PRIOR FILING DATE: 2001-02-08	
		; PRIOR APPLICATION NUMBER: 60/271,021	
		; PRIOR FILING DATE: 2001-02-22	
		; PRIOR APPLICATION NUMBER: 60/275,946	
		; PRIOR FILING DATE: 2001-03-14	
		; PRIOR APPLICATION NUMBER: 60/278,150	
		; PRIOR FILING DATE: 2001-03-23	
		; PRIOR APPLICATION NUMBER: 60/285,718	
		; PRIOR FILING DATE: 2001-04-23	
		; PRIOR APPLICATION NUMBER: 60/312,902	
		; PRIOR FILING DATE: 2001-08-16	
		; PRIOR APPLICATION NUMBER: 60/257,876	
		; PRIOR FILING DATE: 2000-12-21	
		; PRIOR APPLICATION NUMBER: 60/260,718	
		; PRIOR FILING DATE: 2001-01-10	
		; PRIOR APPLICATION NUMBER: 60/284,591	
		; PRIOR FILING DATE: 2001-04-18	
		; NUMBER OF SEQ ID NOS: 352	
		; SOFTWARE: PatentIn Ver. 2.1	
		; SEQ ID NO 35	
		; LENGTH: 998	
		; TYPE: DNA	
		; ORGANISM: Homo sapiens	
		; FEATURE:	
		; NAME/KEY: CDS	
		; LOCATION: (2)..(988)	
		US-10-025-806-35	
		Query Match	
		99.4%; Score 980.6; DB 15; Length 998;	

1 ATGCCAGCACTCTTGGCCACAAACATGAAATCTCTCATCACTGATGTTGACCTTCT 60		1 ATGCCAGCACTCTTGGCCACAAACATGAAATCTCTCATCACTGATGTTGACCTTCT 60	
61 GTCCTTCTCTCTGGGATCCAGGCTGGAACAAATTTCTGCTCTCTCTCTCTCT 120		61 GTCCTTCTCTCTGGGATCCAGGCTGGAACAAATTTCTGCTCTCTCTCTCTCT 120	
121 GTGTGTGCTTAGGCACAGCCCAATTTGGGCAATATAATTTCTGCTGTTGTTGCC 180		121 GTGTGTGCTTAGGCACAGCCCAATTTGGGCAATATAATTTCTGCTGTTGTTGCC 180	
181 ACTGAACAGCTTTGACAAAGCTGTGTACCTTTTCTGTGCATGCTCTCAACCATGCAC 240		181 ACTGAACAGCTTTGACAAAGCTGTGTACCTTTTCTGTGCATGCTCTCAACCATGCAC 240	
241 TTGGCTGCTCTGTCTCCACAGTTCGCAAGCTACTGGCTATCTTCTGTGTGGAGCCGA 300		241 TTGGCTGCTCTGTCTCCACAGTTCGCAAGCTACTGGCTATCTTCTGTGTGGAGCCGA 300	
301 CATATATCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 360		301 CATATATCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 360	
361 GAGTCCACTGTCTACTGCGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420		361 GAGTCCACTGTCTACTGCGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420	
421 CGCTATGCCAATCTCTACTGACACCATCATTTGCCACATAGGGGTGGCAGCTGTAG 480		421 CGCTATGCCAATCTCTACTGACACCATCATTTGCCACATAGGGGTGGCAGCTGTAG 480	
481 CGAGCTCCCTGCTCTATGCTCCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540		481 CGAGCTCCCTGCTCTATGCTCCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540	
541 AGCCATGTGATCCTACACAGCTACTGTGAGCAGCATGGCTGTGGTGAAGCTGGCCTGTGA 600		541 AGCCATGTGATCCTACACAGCTACTGTGAGCAGCATGGCTGTGGTGAAGCTGGCCTGTGA 600	
601 GACACAGGCTTAACGGTGTGTATGGCTGACAGCTGCACTGTTGCTCATTTGGGGTTGAC 660		601 GACACAGGCTTAACGGTGTGTATGGCTGACAGCTGCACTGTTGCTCATTTGGGGTTGAC 660	
661 TTGTTTGTGATTTGCTCTCTATGCTGCTTAATTTGCAAGCTGCTCTCGCTCTCATCC 720		661 TTGTTTGTGATTTGCTCTCTATGCTGCTTAATTTGCAAGCTGCTCTCGCTCTCATCC 720	
721 CATGAAGCTCGGTCCAGGCTTGGGATCTGTTGCTGCTGCTGCTGCTGCTGCTGCTG 780		721 CATGAAGCTCGGTCCAGGCTTGGGATCTGTTGCTGCTGCTGCTGCTGCTGCTGCTG 780	
781 TCTTATACAGGCT 840		781 TCTTATACAGGCT 840	
841 CATATTACATTTCTTTTGGCCAAATGTTTATCTCTCTTTTGGCCACCTGCTCTTAATCTGTG 900		841 CATATTACATTTCTTTTGGCCAAATGTTTATCTCTCTTTTGGCCACCTGCTCTTAATCTGTG 900	
901 GTATATGGAGTTAAGACCAACAGATCCGTAAGAGTTGTTCAGGGTGTTCAGAGTGGG 960		901 GTATATGGAGTTAAGACCAACAGATCCGTAAGAGTTGTTCAGGGTGTTCAGAGTGGG 960	
961 CAGGGAATGGGCATCAAGCATCTGAG 987		961 CAGGGAATGGGCATCAAGCATCTGAG 987	

Best Local Similarity 99.6%; Pred. No. 2.3e-105;			
Matches 983; Conservative 0; Mismatches 4; Indels 0; Gaps 0;			
QY	1	ATGTCAGACACTCTTGGCCACAACTGAATCTCTCATCACACTGATGTTGACCTCTCT	60
Db	2	ATGTCAGACACTCTTGGCCACAACTGAATCTCTCATCACACTGATGTTGACCTCTCT	61
QY	61	GTCTTCTTCTCTCTGGGATCCAGCTCTGGAACAATTTCAATTTGGCTCTCATCTCCCT	120
Db	62	GTCTTCTTCTCTCTGGGATCCAGCTCTGGAACAATTTCAATTTGGCTCTCATCTCCCT	121
QY	121	GTGTGTGGCTTAGGCACAGCCCAATTTGTGGCAATATAAATTAATTTCTGTCTTGGC	180
Db	122	GTGTGTGGCTTAGGCACAGCCCAATTTGTGGCAATATAAATTAATTTCTGTCTTGGC	181
QY	181	ACTGAACACAGCTTTCGACAGCCCTGTACCTTTTCTGTGATGCTCTCAACCATCGAC	240
Db	182	ACTGAACACAGCTTTCGACAGCCCTGTACCTTTTCTGTGATGCTCTCAACCATCGAC	241
QY	241	TTGGCTGCTCTGTCTCCACAGTTCCTCAAGCTACTGCTATCTTCTGTGTGGAGCCGA	300
Db	242	TTGGCTGCTCTGTCTCCACAGTTCCTCAAGCTACTGCTATCTTCTGTGTGGAGCCGA	301
QY	301	CATATATCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	360
Db	302	CATATATCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	361
QY	361	GAGTCCACTGTCTACTGGCCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	420
Db	362	GAGTCCACTGTCTACTGGCCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	421
QY	421	CGCTATGCGCAATCTCACTGACACCATATGCTGCTGCTGCTGCTGCTGCTGCTGCTG	480
Db	422	CGCTATGCGCAATCTCACTGACACCATATGCTGCTGCTGCTGCTGCTGCTGCTGCTG	481
QY	481	CGAGGCTCCCTGCTCATGCTCCCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	540
Db	482	CGAGGCTCCCTGCTCATGCTCCCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	541
QY	541	AGCATATGATCTTACACAGCTACTGAGACATGCTGCTGCTGCTGCTGCTGCTGCTGCT	600
Db	542	AGCATATGATCTTACACAGCTACTGAGACATGCTGCTGCTGCTGCTGCTGCTGCTGCT	601
QY	601	GACACAGGCTTAACCGCTGTATGGCTGACAGCTGCTGCTGCTGCTGCTGCTGCTGCTG	660
Db	602	GACACAGGCTTAACCGCTGTATGGCTGACAGCTGCTGCTGCTGCTGCTGCTGCTGCTG	661
QY	661	TTGTTTTGCAATGGTCTCTCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	720
Db	662	TTGTTTTGCAATGGTCTCTCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	721
QY	721	CATGAAGCTGGTCAAGGCTTACAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	780
Db	722	CATGAAGCTGGTCAAGGCTTACAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	781
QY	781	TCTTATACACAGGCTTC	840
Db	782	TCTTATACACAGGCTTC	841
QY	841	CATATTCACATCTTTTGGCCAAATGTTTATCTGCTTTTGGCCAAATGTTTATCTGCTT	900
Db	842	CATATTCACATCTTTTGGCCAAATGTTTATCTGCTTTTGGCCAAATGTTTATCTGCTT	901
QY	901	GTATATGAGTTAGACAAACAGATCCGTTAAAGAGTTGTGAGGTTGTTTCAAGTGGG	960
Db	902	GTATATGAGTTAGACAAACAGATCCGTTAAAGAGTTGTGAGGTTGTTTCAAGTGGG	961
QY	961	CAGGGAATGGGCATCAAGGCATCTGAG 987	
Db	962	CAGGGAATGGGCATCAAGGCATCTGAG 988	

RESULT 6

US-10-025-806-37
; Sequence 37, Application US/10025806
; Publication No. US20030198955A1
; GENERAL INFORMATION:
; APPLICANT: Li, Li
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Ballinger, Robert
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Colman, Steven
; APPLICANT: Spytek, Kimberly
; APPLICANT: Casman, Stacie
; APPLICANT: Edinger, Shlomit
; APPLICANT: Gerlach, Valerie
; APPLICANT: Sciore, Paul
; APPLICANT: Smithson, Glennda
; APPLICANT: Peyman, John
; APPLICANT: MacDougall, John
; APPLICANT: Stone, David
; APPLICANT: Vernet, Corine
; APPLICANT: Shenoy, Suresh
; APPLICANT: Gunther, Erik
; APPLICANT: Millet, Isabelle
; APPLICANT: Tchernev, Velizar
; APPLICANT: Anderson, David
; APPLICANT: Gusev, Vladimir
; APPLICANT: Malyankar, Uriel
; APPLICANT: Zhong, Haihong
; APPLICANT: Ellerman, Karen
; APPLICANT: Wolenc, Adam
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 21402-224 AB
; CURRENT APPLICATION NUMBER: US/10/025.806
; CURRENT FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 60/256,635
; PRIOR FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: 60/259,743
; PRIOR FILING DATE: 2001-01-04
; PRIOR APPLICATION NUMBER: 60/299,327
; PRIOR FILING DATE: 2001-06-19
; PRIOR APPLICATION NUMBER: 60/261,498
; PRIOR FILING DATE: 2001-01-12
; PRIOR APPLICATION NUMBER: 60/263,689
; PRIOR FILING DATE: 2001-01-24
; PRIOR APPLICATION NUMBER: 60/276,464
; PRIOR FILING DATE: 2001-02-08
; PRIOR APPLICATION NUMBER: 60/271,021
; PRIOR FILING DATE: 2001-02-22
; PRIOR APPLICATION NUMBER: 60/275,946
; PRIOR FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: 60/278,150
; PRIOR FILING DATE: 2001-03-23
; PRIOR APPLICATION NUMBER: 60/285,718
; PRIOR FILING DATE: 2001-04-23
; PRIOR APPLICATION NUMBER: 60/312,902
; PRIOR FILING DATE: 2001-08-16
; PRIOR APPLICATION NUMBER: 60/257,876
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/260,718
; PRIOR FILING DATE: 2001-01-10
; PRIOR APPLICATION NUMBER: 60/284,591
; NUMBER OF SEQ ID NOS: 352
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 37
; LENGTH: 998
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (2)..(988)
US-10-025-806-37

Query Match 99.4%; Score 980.6; DB 15; Length 998;

Db 777 GACACAGCCCTAACCGTGTGTATGGGCTGACAGCTGACATTTGTTGGGTTGAC 836
Qy 661 TTGTTTTCATGTCCTCTATGCCCTAAATGCAAGCTGTCCTTCGCTCTCATCC 720
Db 837 TTGTTTTCATGTCCTCTATGCCCTAAATGCAAGCTGTCCTTCGCTCTCATCC 896
Qy 721 CATGAAGCTCGGTCCAGGCCCTAGGACCTGTTGGTTCCTGCTGTCATCTCATC 780
Db 897 CATGAAGCTCGGTCCAGGCCCTAGGACCTGTTGGTTCCTGCTGTCATCTCATC 956
Qy 781 TCTTATACACAGCCCTCTTCTCTCTCTTTTACACACCCCTTTGGCCATCAGCTTCCAGTC 840
Db 957 TCTTATACACAGCCCTCTTCTCTCTTTTACACACCCCTTTGGCCATCAGCTTCCAGTC 1016
Qy 841 CATATTCACATCTTTTGGCCAAATGTTTATCTGCTTTTGGCCACCTGCTCTTAACTCTGTG 900
Db 1017 CATATTCACATCTTTTGGCCAAATGTTTATCTGCTTTTGGCCACCTGCTCTTAACTCTGTG 1076
Qy 901 GTATATGAGTTAAGACCAACACATCCGTAAAGAGTTGTCAGGGTGTTCCTTCAAGTGGG 960
Db 1077 GTATATGAGTTAAGACCAACACATCCGTAAAGAGTTGTCAGGGTGTTCCTTCAAGTGGG 1136
Qy 961 CAGGGAATGGGCATCAAGGCATCTGAG 987
Db 1137 CAGGGAATGGGCATCAAGGCATCTGAG 1163

RESULT 8

US-10-292-798-797
; Sequence 797, Application US/10292798
; Publication No. US2003023583A1
; GENERAL INFORMATION:
; APPLICANT: SUMA, MAKIKO
; APPLICANT: ASAI, KIYOSHI
; APPLICANT: AKIYAMA, YUTAKA
; APPLICANT: ABURATANI, HIROYUKI
; TITLE OF INVENTION: GUANOSINE TRIPHOSPHATE-BINDING PROTEIN COUPLED RECEPTORS
; FILE REFERENCE: 084335/166
; CURRENT APPLICATION NUMBER: US/10/292,798
; CURRENT FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: 10/017,161
; PRIOR FILING DATE: 2001-12-18
; PRIOR APPLICATION NUMBER: JP 2001-246789
; PRIOR FILING DATE: 2001-06-18
; NUMBER OF SEQ ID NOS: 2070
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 797
; LENGTH: 1390
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; LOCATION: source
; FEATURE:
; LOCATION: (1)..(1390)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (201)..(1190)
US-10-292-798-797

Query Match 99.4%; Score 980.6; DB 16; Length 1390;
Best Local Similarity 99.6%; Pred. No. 2.8e-305;
Matches 983; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
Qy 1 ATGTCCAGCACTCTTGGCCCAACATGGAATCTCTCATCACACTGATGTTGACCCCTTCT 60
Db 201 ATGTCCAGCACTCTTGGCCCAACATGGAATCTCTCATCACACTGATGTTGACCCCTTCT 260
Qy 61 GTCTTCTCTCTCCCTGGGCATCCAGGCTCGGAACAATTCATTTGGTGGCTCTCATCCCT 120
Db 261 GTCTTCTCTCTCCCTGGGCATCCAGGCTCGGAACAATTCATTTGGTGGCTCTCATCCCT 320
Qy 121 GTGTGTGGCTTAGGCACAGCCACAATTTGTGGGCAATATACTATTCTCTGTTGTTGGCC 180

Db 321 GTGTGTGGCTTAGGCACAGCCACAATTTGTGGGCAATATACTATTCTCTGTTGTTGGCC 380
Qy 181 ACTGAACCCAGTCTTGGCAACAGCCTGTGTACCTTTTCTGTGCATGCTCTCAACCATCCAC 240
Db 381 ACTGAACCCAGTCTTGGCAACAGCCTGTGTACCTTTTCTGTGCATGCTCTCAACCATCCAC 440
Qy 241 TTGCTGTGCTCTGTCTCCACAGTTTCCCAAGTACTTGGCTATCTTCTGTGTGTGGAGCCGGA 300
Db 441 TTGCTGTGCTCTGTCTCCACAGTTTCCCAAGTACTTGGCTATCTTCTGTGTGTGGAGCCGGA 500
Qy 301 CATATATCTGCTCTGCTGCTGCGCACATATGTTCTTCTTATTCATGCTCTTGTGATGATG 360
Db 501 CATATATCTGCTCTGCTGCTGCGCACATATGTTCTTCTTATTCATGCTCTTGTGATGATG 560
Qy 361 GAGTCCACTGTGCTACTTGGCCATGGCCCTTTGATCGCTACGTGGCCATCTGCCACCCACTC 420
Db 561 GAGTCCACTGTGCTACTTGGCCATGGCCCTTTGATCGTACGTGGCCATCTGCCACCCACTC 620
Qy 421 CGCTATGCGCACATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 480
Db 621 CGCTATGCGCACATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 680
Qy 481 CGAGGCTCCCTGCTCATGCTCCCATGCTCCCTTCTTATTTGGGCGTTTGAACCTTCTGCCAA 540
Db 581 CGAGGCTCCCTGCTCATGCTCCCATGCTCCCTTCTTATTTGGGCGTTTGAACCTTCTGCCAA 740
Qy 541 AGCCATGTGATCTCTACACAGCTACTGTGAGCACATGGCTGTGTGAAAGCTGGCCTGTGGA 600
Db 741 AGCCATGTGATCTCTACACAGCTACTGTGAGCACATGGCTGTGTGAAAGCTGGCCTGTGGA 800
Qy 601 GACACAGGCTTAACCGTGTGTATGGGCTGACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660
Db 801 GACACAGGCTTAACCGTGTGTATGGGCTGACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 860
Qy 661 TTGTTTGTGCTGCTCTCTCTATGCTTCTCTTATTCACAAAGCTGCTTTCGCTCTCTCATCC 720
Db 861 TTGTTTGTGCTGCTCTCTCTATGCTTCTCTTATGCTTCTTATGCTTCTTCTGCTCTCTCATCC 920
Qy 721 CATGAAGCTCGGTCCAAAGGCCCTAGGACCTGTGGTTCCTCATGTCTGTGCTCATCTCATC 780
Db 921 CATGAAGCTCGGTCCAAAGGCCCTAGGACCTGTGGTTCCTCATGTCTGTGCTCATCTCATC 980
Qy 781 TCTTATACACAGCCCTCTTCTCTCTTTTACACACCGCTTTGGCCATCAGCTTCCAGTC 840
Db 981 TCTTATACACAGCCCTCTTCTCTCTTTTACACACCGCTTTGGCCATCAGCTTCCAGTC 1040
Qy 841 CATATTCACATCTTTTGGCCAAATGTTTATCTGCTTTTGGCCACCTGCTTCTTAACTCTGTG 900
Db 1041 CATATTCACATCTTTTGGCCAAATGTTTATCTGCTTTTGGCCACCTGCTTCTTAACTCTGTG 1100
Qy 901 GTATATGAGTTAAGACCAACACATCCGTAAAGAGTTGTCAGGGTGTTCCTTCAAGTGGG 960
Db 1101 GTATATGAGTTAAGACCAACACATCCGTAAAGAGTTGTCAGGGTGTTCCTTCAAGTGGG 1160
Qy 961 CAGGGAATGGGCATCAAGGCATCTGAG 987
Db 1161 CAGGGAATGGGCATCAAGGCATCTGAG 1187

RESULT 9

US-10-297-021-42
; Sequence 42, Application US/10297021
; Publication No. US20040023294A1
; GENERAL INFORMATION:
; APPLICANT: INCYTE GENOMICS, INC.
; APPLICANT: PATTERSON, Chandra
; APPLICANT: TRIBOULEY, Catherine M.
; APPLICANT: YAO, Monique G.
; APPLICANT: GRIFFIN, Jennifer A.
; APPLICANT: THORNTON, Michael
; APPLICANT: LU, Yan
; APPLICANT: KALLICK, Deborah A.

QY	25	ATGGAATCTCTCATCACTGATGTTGACCTTCTGTCTTCTTCTCTCTCTGGCATCCCA	84
Db	1	ATGTGAGCCTCCAATATCACCTTAACACATCCAACTGCTTCTTGTGTGGGATTC	60
QY	85	GGCTGGAACAAATTTCATTTGTGCTCTCACTCCCTGTGTGTGGCTTAGGCACAGCCACA	144
Db	61	GGCTGGAACACCTGCAATCTGATCTCCATCCCTTTCTGTTAGCATATACACTGGCC	120
QY	145	ATTGTGGCAATAAATATATCTGGTTGTTGTGCCACTGAACCACTTGCACAGCCT	204
Db	121	CTGCTTGAAACTGCATCTCCTTCTCATATCCAGCTGATGAGCCCTCCATGAAGCC	180
QY	205	GTGTACCTTTTTCTGTGCATGCTCTCAACCACTCGACTTGGCTGCCTGTCTCCACAGTT	264
Db	181	ATGTACCTCTTTCTGGCCATGTTGGCAGCCATCGACCTGTGTCTTCTCTCAGCATG	240
QY	265	CCCAAGCTACTGGCTATCTTCTGTGTGGAGCCGACATATATCTGCCCTCTGCCTGCCG	324
Db	241	CCCAAAATGCTTGCATATTTCTGGTTCAGGGATCGGGAGATAAACTTCTTTTGGCTGTCTG	300
QY	325	GCATATATGTTCTTCATTCATGCTTCTGCAATGATGGAGTCCACTGCTACTGCCCATG	384
Db	301	GCCAGATGTTCTTCTCTCACTCCTCTCTCCATCATGGAGTCAGCAGTCTGTGGCCCATG	360
QY	385	GCTTTTCATCGTACGTGGCCATCTGCCACCACCTCCGCTATGCCACAAATCCTCACTGAC	444
Db	361	GCTTTTGACCCGCTATGTGGCTATCTGCAAGGCATGCACTACCAAGGTCTGTACTGGG	420
QY	445	ACCATCATTTGCCACATAGGGGTGGCAGCTGTAGTTCGAGGCTCCCTGCTCATGCTCCCA	504
Db	421	TCCCTCATCACAAGATGGCATGGCTGCTGTGGCCGGGTGTGACACTAATGACTCCA	480
QY	505	TGTCCTCTCTTTATTTGGGCGTTTGAACCTTCTGCCAAAGCCATGTGATCCTACACACGTAC	564
Db	481	CTCCCTTCTGCTGAGATGTTTCCACTACTGCCGAGGCCAGTGCCTCACTGCTACT	540
QY	565	TGTGAGCACATGGCTGTGGTGAAGCTGGCCTGTGGAGACAACGAGCCTAAACCGTGTGAT	624

[illegible]

Db 61 GGCTGGAAACACCTGACATCTGGATCTCCATCCCTTCTGCTTAGCATATACACTGGCC 120
Qy 145 ATTGTGGGCAATTAATAATTCTGGTGTGTGGTGGCCACTGAAACAGCTCTTGCAACAAGCCT 204
Db 121 CTGCTTGGAAACTGCACTCTCTCTCTCATATCAGGCTGATGAGCCCTCCATGAACCC 180
Qy 205 GTGTACCTTTTCTGTCATGCTCTCAACCATGACCTGGCTGCTCTGTCTCCACAGTT 264
Db 181 ATGTACCTCTTTCTGGGCAATGTTGGCAGCCATGACCTGGCTCTCTCTCAGCACTG 240
Qy 265 CCCAAGCTACTGGCTATCTCTCTGTGTGGAGCCGACATATATCTGCTCTGCTGCTGCTG 324
Db 241 CCCAATATGCTTGGCAATCTCTGTTGAGGATCGGAGATTAACCTTCTTGTCTGCTGCTG 300
Qy 325 GCACATATGTTCTTCAATCATGCTCTTGTGATGAGTGCATCTGTGCTACTGGCCATG 384
Db 301 GCCCAGATGTTCT 360
Qy 385 GCTTTGATCGTACGTCAGTGGCCATCTGCCACCATCTCGCTATGCCAATCTCTCACTGAC 444
Db 361 GCTTTGACCGCTATGCTGCTATCTGCAAGCACTGCACTACACCAAGTCTCTGACTGGG 420
Qy 445 ACCATCATGCCCCATAGGGGTGGCAGCTGTAGTGGAGGCTCCCTGCTCATCTCCCA 504
Db 421 TCCCTCATCACAAGATTGGATGGCTGCTGTGGCCCGGGCTGACACTAATGACTCCA 480
Qy 505 TGTCCCTCTTTATTGGCGTTTGAACCTTCTGCGCAAGCCATGATGCTTACACAGCTAC 564
Db 481 CTCCCT 540
Qy 565 TGTGACACATGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 624
Db 541 TGTGAACACATGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 600
Qy 625 GGGCTGACAGCTGACCTGTTGGCTCAITGGGGTGTGACTTGTGTTGCAATGGTCTCTCT 684
Db 601 GGCAATGCTGTGGCAATGTTATTGTTGGTGTGGCTGCTGCTGCTGCTGCTGCTGCTGCT 660
Qy 685 GCCCTAATGCAAGCTGCTCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 744
Db 661 ATCTTTATTCTCAGGAGTCT 720
Qy 745 GGGACCTGTGTTTCCCATGCTGTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 804
Db 721 GGGACATGCTGCT 780
Qy 805 TTTTTCACACCGCTTTGGCCATCAGCTTCCAGTCCATATTCATCTTTTGGCCAAAT 864
Db 781 TCAGTCAATGAGATTGTGAGGGTGTTCAGAG 956
Qy 925 ATCCGTAAGAGATTGTGAGGGTGTTCAGAG 956
Db 901 ATCCGTGAGACATCTTGGAGATATCCCAAG 932

RESULT 15

US-10-182-822A-25

; Sequence 25, Application US/10182822A

; Publication No. US20030211493A1

; GENERAL INFORMATION:

; APPLICANT: INCYTE GENOMICS, INC.; BAUGHN, Mariah R.

; APPLICANT: AU-YOUNG, Janice; YUE, Henry

; TITLE OF INVENTION: G-PROTEIN COUPLED RECEPTORS

; FILE REFERENCE: PI-0032 USN

; CURRENT APPLICATION NUMBER: US/10/182,822A

; CURRENT FILING DATE: 2001-02-01

; PRIOR APPLICATION NUMBER: PCT/US 01/03455

; PRIOR FILING DATE: 2001-01-02

; PRIOR APPLICATION NUMBER: US 60/180,093

; PRIOR FILING DATE: 2000-02-02
; PRIOR APPLICATION NUMBER: US 60/182,045
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: PERL Program
; SEQ ID NO 25
; LENGTH: 945
; TYPE: DNA
; ORGANISM: Homo sapiens
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No: 7472036CB1
US-10-182-822A-25

Query Match 38.0%; Score 375.2; DB 13; Length 945;

Best Local Similarity 62.7%; Pred. No. 5.6e-110;

Matches 584; Conservative 0; Mismatches 348; Indels 0; Gaps 0;

Qy 25 ATGGAATCTCTCATACACTGATGTTGACCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 84
Db 1 ATGTGACCTTCCAAATATACCTTAAACATCCCACTGCTTCTTGTGGTGGGATTTCCA 60
Qy 85 GGTCTGGAACAATTTCAATTTTGTGCTCTCACTCCCTGTGTGTGGCTTAGGCACAGCCACA 144
Db 61 GGCCTGGAACAACCTGCACTCTGGATCTCCATCCCTTCTGCTTAGCATATACACTGGCC 120
Qy 145 ATTGTGGGCAATTAATAATTCTGGTGTGTGGTGGCCACTGAAACAGCTCTTGCAACAAGCCT 204
Db 121 CTGCTTGGAAACTGCACTCTCTCTCTCATATCAGGCTGATGAGCCCTCCATGAACCC 180
Qy 205 GTGTACCTTTTCTGTCATGCTCTCAACCATGACCTTGGCTGCTCTGTCTCCACAGTT 264
Db 181 ATGTACCTCTTTCTGGGCAATGTTGGCAGCCATGCACTGGTCTCTCTCTCAGCACTG 240
Qy 265 CCCAAGCTACTGGCTATCTCTCTGTGTGGAGCCGACATATATCTGCTCTGCTGCTGCTG 324
Db 241 CCCAATATGCTTGGCAATCTCTGTTGAGGATCGGAGATTAACCTTCTTGTCTGCTGCTG 300
Qy 325 GCACATATGTTCTTCAATCATGCTCTTGTGATGAGTGCATCTGTGCTACTGGCCATG 384
Db 301 GGGCAGATGTTCT 360
Qy 385 GCTTTGATCGTACGTCAGTGGCCATCTGCCACCATCTCGCTATGCCAATCTCTCACTGAC 444
Db 361 GCTTTGACCGCTATGCTGCTATCTGCAAGCACTGCACTACACCAAGTCTCTGACTGGG 420
Qy 445 ACCATCATGCCCCATAGGGGTGGCAGCTGTAGTGGAGGCTCCCTGCTCATGCTCCCA 504
Db 421 TCCCTCATCACAAGATTGGCAATCTCTGTTGAGGATCGGAGATTAACCTTCTTGTCTGCTGCT 480
Qy 505 TGTCCCTCTTTATTGGGCGTTTGAACCTTCTGCGCAAGCCATGATGCTTACACAGCTAC 564
Db 481 CTCCCT 540
Qy 565 TGTGACACATGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 624
Db 541 TGTGAACACATGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 600
Qy 625 GGGCTGACAGCTGACCTGTTGGCTCAITGGGGTGTGACTTGTGTTGCAATGGTCTCTCT 684
Db 601 GGCAATGCTGTGGCAATGTTATTGTTGGTGTGGCTGCTGCTGCTGCTGCTGCTGCTGCT 660
Qy 685 GCCCTAATGCAAGCTGCTCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 744
Db 661 ATCTTTATTCTCAGGAGTCT 720
Qy 745 GGGACCTGTGTTTCCCATGCTGTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 804
Db 721 GGGACATGCTGCT 780
Qy 805 TTTTTCACACCGCTTTGGCCATCAGCTTCCAGTCCATATTCATCTTTTGGCCAAAT 864
Db 781 TCAGTCAATGAGATTGTGAGGGTGTTCAGAG 956
Qy 925 ATCCGTAAGAGATTGTGAGGGTGTTCAGAG 956
Db 901 ATCCGTGAGACATCTTGGAGATATCCCAAG 932

Qy	865	GTATTATCTGCTTTTGCCACCTGCTCTTAATCCTGTGGTATATGGAGTTAAGACCAACAG	924
Db	841	TTCTATCTGCTCTTCCCACTGCTCAATCCCAATCTATGSGTCAAGACCAAGCAA	900
Qy	925	ATCCGTAAAGAGTTGTGAGGGTGTTCAAAG	956
Db	901	ATCCGTGAGAGCATCTTGGGAGTATCCCAAG	932

Search completed: August 27, 2004, 18:39:24
 Job time : 544 secs

This Page Blank (upside)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 27, 2004, 15:14:54 ; Search time 4169 Seconds
(without alignments)
10261.352 Million cell updates/sec

Title: US-10-081-775-1_COPY_537_1523

Perfect score: 987

Sequence: 1 atgtccagcactctggcca.....tgggcatacagcatctgag 987

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*

1: gb.ba.*
2: gb.htg.*
3: gb.in.*
4: gb.om.*
5: gb.ov.*
6: gb.pat.*
7: gb.ph.*
8: gb.pl.*
9: gb.pr.*
10: gb.ro.*
11: gb.sts.*
12: gb.sy.*
13: gb.un.*
14: gb.vi.*
15: em.ba.*
16: em.fun.*
17: em.hum.*
18: em.in.*
19: em.mu.*
20: em.on.*
21: em.or.*
22: em.ov.*
23: em.pat.*
24: em.ph.*
25: em.pl.*
26: em.ro.*
27: em.sts.*
28: em.un.*
29: em.vi.*
30: em.htg.hum.*
31: em.htg.inv.*
32: em.htg.other.*
33: em.htg.mus.*
34: em.htg.pln.*
35: em.htg.rod.*
36: em.htg.mam.*
37: em.htg.vit.*
38: em.sv.*
39: em.htgo.hum.*
40: em.htgo.mus.*
41: em.htgo.other.*

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	983.8	99.7	185330	2	AC124272 Homo sapi
2	983.8	99.7	203352	2	AC111179 Homo sapi
3	980.6	99.4	990	6	BD144452 Novel G-p
4	980.6	99.4	998	6	AX551417 Sequence
5	980.6	99.4	998	6	AX551419 Sequence
6	980.6	99.4	1390	6	AX646605 Sequence
7	980.6	99.4	1390	9	AB065534 Homo sapi
8	980.6	99.4	172027	2	AC026090 Homo sapi
9	980.6	99.4	176597	9	AC116156 Homo sapi
10	956.6	96.9	963	6	AX241675 Sequence
11	956.6	96.9	966	6	AX318261 Sequence
12	956.6	96.9	966	6	AX448757 Sequence
13	953.2	96.6	202761	9	AC109341 Homo sapi
14	822.2	83.3	181386	2	AC147586 Ocolemur
15	822.2	83.3	239335	2	AC146635 Ocolemur
16	791	80.1	948	6	AX659373 Sequence
17	761.4	77.1	264026	2	AC113720 Rattus no
18	759.8	77.0	990	10	AY073012 Mus muscu
19	759.8	77.0	191432	2	AC135109 Mus muscu
20	747	75.7	966	10	AY317777 Mus muscu
21	629	63.7	781	4	AY355971 Canis fam
22	376.2	38.1	980	6	AX320273 Sequence
23	376.2	38.1	148313	9	AC010930 Homo sapi
24	376.2	38.1	170338	2	AC138149 Homo sapi
25	375.2	38.0	942	6	AX241840 Sequence
26	375.2	38.0	945	6	AX207657 Sequence
27	375.2	38.0	945	6	AX686665 Sequence
28	375.2	38.0	945	6	BD144356 Novel G-p
29	374.6	38.0	1345	6	AX646565 Sequence
30	374.6	38.0	1345	9	AB065791 Homo sapi
31	374.6	38.0	141485	9	AC009758 Homo sapi
32	373.6	37.9	945	6	AX448711 Sequence
33	368.4	37.3	1013	6	AX551407 Sequence
34	368.4	37.3	1360	6	AX446829 Sequence
35	368.4	37.3	1360	9	AB065820 Homo sapi
36	368.4	37.3	165682	9	AC111177 Homo sapi
37	368.4	37.3	165729	2	AC009642 Homo sapi
38	367.2	37.2	982	6	AX320269 Sequence
39	367.2	37.2	235030	2	AC107331 Rattus no
40	365	37.0	1345	6	AX646563 Sequence
41	365	37.0	1345	9	AB065790 Homo sapi
42	364	36.9	942	6	AX241838 Sequence
43	364	36.9	945	6	AX448713 Sequence
44	364	36.9	1151	6	AX320267 Sequence
45	362.4	36.7	945	6	BD144355 Novel G-p

ALIGNMENTS

RESULT 1
AC124272
LOCUS Homo sapiens chromosome 11 clone RP11-607K3 map 11, WORKING DRAFT
DEFINITION AC124272
ACCESSION AC124272.2 GI:22857733
VERSION HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 185330)
AUTHORS Birren,B., Nusbaum,C. and Lander,E.
TITLE Homo sapiens chromosome 11, clone RP11-607K3

Pred. No. is the number of results predicted by chance to have a

JOURNAL
REFERENCE
AUTHORS

Unpublished

2 (bases 1 to 185330)

Birren, B., Linton, L., Nusbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S., Barna, N., Bastien, V., Bloom, T., Boguslavsky, L., Boukhgalter, B., Brown, A., Camarata, J., Campotiano, A., Chang, J., Chazaro, B., Choepel, Y., Colangelo, M., Collins, S., Collymore, A., Cook, A., Cooke, P., DeArellano, K., Dewar, K., Diaz, J. S., Dodge, S., Faro, S., Ferreira, P., Fitzgerald, M., Fitzhugh, W., Gage, D., Galagan, J., Gardnya, S., Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N., Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Katat, A., Karatas, A., Kells, C., LaRoquette, K., Lamazares, R., Landers, T., Lehoczy, J., Levine, R., Lindblad-Toh, K., Liu, G., MacLean, C., MacDonald, P., Major, J., Marquis, N., Matthews, C., McCarthy, M., McSwan, P., McKernan, K., Meldrim, J., Menseus, L., Mihova, T., Mienga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., Norman, C. H., O'Connor, T., O'Donnell, P., O'Neil, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V., Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P., Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schupback, R., Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N., Strauss, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Travis, N., Triggilio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

TITLE	JOURNAL	REFERENCE	AUTHORS
1. The Effect of Temperature on the Rate of Reaction of Hydrogen Peroxide with Potassium Iodide	Journal of Chemical Education	1985	Smith, J. D.
2. The Effect of Concentration on the Rate of Reaction of Hydrogen Peroxide with Potassium Iodide	Journal of Chemical Education	1985	Smith, J. D.
3. The Effect of pH on the Rate of Reaction of Hydrogen Peroxide with Potassium Iodide	Journal of Chemical Education	1985	Smith, J. D.
4. The Effect of Catalyst on the Rate of Reaction of Hydrogen Peroxide with Potassium Iodide	Journal of Chemical Education	1985	Smith, J. D.
5. The Effect of Surface Area on the Rate of Reaction of Hydrogen Peroxide with Potassium Iodide	Journal of Chemical Education	1985	Smith, J. D.

```
misc_feature
1. .l20991
/note="assembly_fragment
clone_end:Sp6
vector_side:left"
l21052. .l32589
/note="assembly_fragment"
l32690. .l53921
/note="assembly_fragment"
l54022. .l85330
/note="assembly_fragment"

ORIGIN
```

Query Match	99.7%;	Score 983.8;	DB 2;	Length 185330;
Best Local Similarity	99.8%;	Pred. No. 2.6e-287;		
Matches 985;	Conservative 0;	Mismatches 2;	Indels 0;	Gaps 0;
QY	1	ATGTCACGACACTCTTGGCCACACATGGGAATCTCCTCATCACACTGATGTGGACCCCTTCT	60	
DB	3299	ATGTCACGACACTCTTGGCCACACATGGGAATCTCCTCATCACACTGATGTGGACCCCTTCT	3358	
QY	61	GTCTTCTTCTCCTCTGGGCATCCCAAGTCTCGGAACAATTTTCATTGTGGCTCTCACTCCCT	120	
DB	3359	GTCTTCTTCTCCTCTGGGCATCCCAAGTCTCGGAACAATTTTCATTGTGGCTCTCACTCCCT	3416	
QY	121	GTGTGTGGCTTAGGCACAGCCACAAATGTGGGCAATATAACTATTCTGGTGTGTGTGGCC	180	
DB	3419	GTGTGTGGCTTAGGCACAGCCACAAATGTGGGCAATATAACTATTCTGGTGTGTGTGGCC	3476	
QY	181	ACTGAACCAAGCTTTGCACAAAGCCTGTGTACCTTTTCTGTGTCATGCTCTCAACCATCGAC	240	
DB	3479	ACTGAACCAAGCTTTGCACAAAGCCTGTGTACCTTTTCTGTGTCATGCTCTCAACCATCGAC	3538	
QY	241	TTGGCTGCCTCTCTCTCCACAGTTCCTCCAAAGCTACTGGCTATCTTCTGGTGTGGAGCCGGA	300	
DB	3539	TTGGCTGCCTCTCTCTCCACAGTTCCTCCAAAGCTACTGGCTATCTTCTGGTGTGGAGCCGGA	3598	
QY	301	CATATATCTGCCTCTGCCTGCCACATATGTTCTTCATTCAATGCTTCTGCATGATG	360	
DB	3599	CATATATCTGCCTCTGCCTGCCACATATGTTCTTCATTCAATGCTTCTGCATGATG	3655	
QY	361	GAGTCACACTGTGCTACTGGCCATGGCCTTTGATCGCTACGTGGCCATCTGCCACCCATCTC	420	
DB	3659	GAGTCACACTGTGCTACTGGCCATGGCCTTTGATCGCTACGTGGCCATCTGCCACCCATCTC	3711	
QY	421	CGCTATGGCACAAATCCTCACTGACACCATCATTTGCCACATAGGGGTGGCAGCTGTAGTG	480	
DB	3719	CGCTATGGCACAAATCCTCACTGACACCATCATTTGCCACATAGGGGTGGCAGCTGTAGTG	3771	
QY	481	CGAGGCTCCCTGCTCATGCTCCCATGTGCCCTTCTTTATTGGGCGGTTTGAACCTTCGCCAA	540	


```
Db 3779 CGAGGCTCCCTCATGCTCCCATGCTCCCTTTATTTGCGGTGTTGACATCTTGCCTAA 3838
QY 541 AGCCATGATCTTACACAGTACTGTGAGCAGATGGCTGTGGTGAAGTGGCCCTGTGGA 600
Db 3839 AGCCATGATCTTACACAGTACTGTGAGCAGATGGCTGTGGTGAAGTGGCCCTGTGGA 3898
QY 601 GACACAGGCTTAACCGTGTGTATGGCTGACAGCTGACATCTGTTGTCATTTGGGCTTGAC 560
Db 3899 GACACAGGCTTAACCGTGTGTATGGCTGACAGCTGACATCTGTTGTCATTTGGGCTTGAC 3958
QY 661 TTGTTTTCATTTGCTCTCTCTATGCTTAAATGACCAAGCTGTCTTCCGCTCTCATCC 720
Db 3959 TTGTTTTCATTTGCTCTCTCTATGCTTAAATGACCAAGCTGTCTTCCGCTCTCATCC 4018
QY 721 CATGAAGCTCGTCAAGGCTTACGAGCTGTGTTTCCATGTGTGTCATCCTCATC 780
Db 4019 CATGAAGCTCGTCAAGGCTTACGAGCTGTGTTTCCATGTGTGTCATCCTCATC 4078
QY 781 TCTTATACACAGGCTCTCTCTCTTTTACACACCGCTTGGCCATCAGCTTCCAGTC 840
Db 4079 TCTTATACACAGGCTCTCTCTCTTTTACACACCGCTTGGCCATCAGCTTCCAGTC 4138
QY 841 CATATTCATTTCTTTGGCCAAATGTTATCTGCTTTTGGCCACCTGCTCTTAATCCTGTG 900
Db 4139 CATATTCATTTCTTTGGCCAAATGTTATCTGCTTTTGGCCACCTGCTCTTAATCCTGTG 4198
QY 901 GTATATGAGTTAGACCAACAGATCGTAAAGAGTGTGTCAGGTTTCAAGTGGG 960
Db 4199 GTATATGAGTTAGACCAACAGATCGTAAAGAGTGTGTCAGGTTTCAAGTGGG 4258
QY 961 CAGGAATGGGCATCAAGGCATCTGAG 987
Db 4259 CAGGAATGGGCATCAAGGCATCTGAG 4285

RESULT 2
AC111179 203352 bp DNA linear HTG 06-AUG-2002
LOCUS Homo sapiens chromosome 11 clone RP11-589G14 map 11, WORKING DRAFT
DEFINITION AC111179
ACCESSION AC111179.2 GI:21699469
VERSION HTG; HTGS PHASE2; HTGS_DRAFT; HTGS_FULLTOP.
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 203352)
AUTHORS Birren, B., Linton, L., Nussbaum, C., Lander, E., Ali, A., Allen, N.,
TITLE Anderson, S., Barna, N., Bastien, V., Boguslavsky, L., Boukhgalter, B.,
JOURNAL Brown, A., Camarata, J., Campiolo, A., Chang, J., Chazaro, B.,
AUTHORS Choepel, Y., Colangelo, M., Collins, S., Collymore, A., Cook, A.,
Cooke, P., DeRellano, K., Dewar, K., Diaz, J.S., Dodge, S., Faro, S.,
Ferreira, P., FitzHugh, W., Gage, D., Galagan, J., Gardyna, S.,
Hagob, S., Gordon, S., Goyette, M., Graham, L., Johnson, R., Jones, C.,
Ginde, B., Horton, L., Hulme, W., Iliiev, I., Johnson, R., Jones, C.,
Kamat, A., Karatas, A., Kelle, C., LaRocque, K., Lamazares, R.,
Landers, T., Leheczy, J., Levine, R., Liu, G., Maclean, C.,
Macdonald, P., Major, J., Marquis, N., Matthews, C., McCarthy, M.,
McKwan, P., McKernan, K., Meldrum, J., Meneus, L., Mihova, T.,
Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C.,
Norman, C.H., O'Connor, T., O'Donnell, P., O'Neil, D., Oliver, J.,
Peterson, K., Phunkhang, P., Pierre, N., Pollara, V., Raymond, C.,
Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P., Roman, J.,
Rosetti, M., Roy, A., Santos, R., Schauer, S., Schupback, R., Seaman, S.,
Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N.,
Strauss, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J.,
Topham, K., Travers, M., Travis, N., Trigilio, J., Vassiliev, H.,
```

```
TITLE
JOURNAL
REFERENCE
AUTHORS
Birren, B., Nussbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S.,
Barna, N., Bastien, V., Bloom, T., Boguslavsky, L., Boukhgalter, B.,
Camarata, J., Chang, J., Chazaro, B., Choepel, Y., Collymore, A.,
Cook, A., Cooke, P., DeRellano, K., Dewar, K., Diaz, J.S., Dodge, S.,
Faro, S., Ferreira, P., FitzGerald, M., Gage, D., Galagan, J.,
Gardyna, S., Gordon, S., Graham, L., Grand-Pierre, N., Hagob, S.,
Horton, L., Hulme, W., Iliiev, I., Johnson, R., Jones, C., Kamat, A.,
Karatas, A., Kelle, C., Landers, T., Levine, R., Lindblad-Toh, K.,
McCarthy, M., Meldrum, J., Meneus, L., Mihova, T., Matthews, C.,
Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., Norman, C.H.,
O'Connor, T., O'Donnell, P., O'Neil, D., Oliver, J., Peterson, K.,
Phunkhang, P., Pierre, N., Raymond, C., Retta, R., Rise, C., Rogov, P.,
Roman, J., Roy, A., Schauer, S., Schupback, R., Seaman, S., Severy, P.,
Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Talamas, J.,
Tesfaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H.,
Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J.,
Zembek, L., Zimmer, A. and Zody, M.
Direct Submission
Submitted (06-AUG-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIDR
Web site: http://www-seq.wi.mit.edu
Contact: sequence.submissions@genome.wi.mit.edu
----- Project Information
Center project name: L25440
Center clone name: 589 G.14
----- Summary Statistics
Sequencing vector: plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 201894 bases at least Q40
Consensus quality: 202347 bases at least Q30
Consensus quality: 202672 bases at least Q20
Insert size: 200000; agarose-fp
Insert size: 203052; sum-of-contigs
Quality coverage: 14.6 in Q20 bases; agarose-fp
Quality coverage: 14.4 in Q20 bases; sum-of-contigs
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 4 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* been provided by the submitter.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
* 1 70426: contig of 70426 bp in length
* 70427 70526: gap of 100 bp
* 70527 78730: contig of 8204 bp in length
* 78731 78830: gap of 100 bp
* 78831 137440: contig of 58610 bp in length
* 137441 137540: gap of 100 bp
* 137541 203352: contig of 65812 bp in length.
Location/Qualifiers
1. 203352
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon,9606"
/chromosome="11"
FEATURES
source
```

```
/map="11"
/clone="RP11-589G14"
/clone.lib="RP11-II Human Male BAC"
1..70426
/misc_feature
  clone_end:sp6
  vector_side:left
70527..78730
/misc_feature
  /note="assembly_fragment"
78831..137440
/misc_feature
  /note="assembly_fragment"
137541..203352
/misc_feature
  /note="assembly_fragment"
  clone_end:77
  vector_side:right

Query Match      99.7%; Score 983.8; DB 2; Length 203352;
Best Local Similarity 99.8%; Pred. No. 2.7e-287;
Matches 985; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGTCACGACTTTGGCCACAACATGGAATCTCTCATCACTGATGTTGACCTTCT 60
DB 169277 ATGTCACGACTTTGGCCACAACATGGAATCTCTCATCACTGATGTTGACCTTCT 169336
QY 61 GTCTTCTTCTCTGGGATCCAGGCTGGNACAAATTTTCTTTGGCTCTCACTCCCT 120
DB 169337 GTCTTCTTCTCTGGGATCCAGGCTGGNACAAATTTTCTTTGGCTCTCACTCCCT 169396
QY 121 GTGTGTGCTTAGGCACAGCACAAATCTGGGCAATATAACTATTTCTGTTGTGTC 180
DB 169397 GTGTGTGCTTAGGCACAGCACAAATCTGGGCAATATAACTATTTCTGTTGTGTC 169456
QY 181 ACTGAACAGCTTTCACAAGCCTGTGACCTTTTCTGTGCATGCTCTCAACCATCGAC 240
DB 169457 ACTGAACAGCTTTCACAAGCCTGTGACCTTTTCTGTGCATGCTCTCAACCATCGAC 169516
QY 241 TTGGCTGCTCTGTCTCCACAGTTCCTCAAGCTACTGGCTATCTTCTGGTGGAGCCGA 300
DB 169517 TTGGCTGCTCTGTCTCCACAGTTCCTCAAGCTACTGGCTATCTTCTGGTGGAGCCGA 169576
QY 301 CATATATCTGCTCTGCTGCTGCGCACATATGTTCTTCAATCATGCTCTTGCATGATG 360
DB 169577 CATATATCTGCTCTGCTGCTGCGCACATATGTTCTTCAATCATGCTCTTGCATGATG 169636
QY 361 GAGTCCACTGTGCTACTGCGCAATGCGCTTTGATCGCTAGTGGCCATCTGCCACCACTC 420
DB 169637 GAGTCCACTGTGCTACTGCGCAATGCGCTTTGATCGCTAGTGGCCATCTGCCACCACTC 169696
QY 421 CGCTATGCCAATCTCTCACTGACACCATCATTTGCCACATAGGCGTGGCAGCTGTAGTG 480
DB 169697 CGCTATGCCAATCTCTCACTGACACCATCATTTGCCACATAGGCGTGGCAGCTGTAGTG 169756
QY 481 CGAGGCTCCCTGCTCATGCTCCCATGTCCTCTTTTATTTGGCGTTTGAACCTCTGCCAA 540
DB 169757 CGAGGCTCCCTGCTCATGCTCCCATGTCCTCTTTTATTTGGCGTTTGAACCTCTGCCAA 169816
QY 541 AGCCATGTATGCTTACACAGTACTGTGTGAGCAATGCTGTGTGTGTGTGTGTGTGTGA 600
DB 169817 AGCCATGTATGCTTACACAGTACTGTGTGAGCAATGCTGTGTGTGTGTGTGTGTGA 169876
QY 601 GACACAGGCTTAACCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGAC 660
DB 169877 GACACAGGCTTAACCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGAC 169936
QY 661 TTGTTTTGCAATGGTCTCTCTATGCTTATGCTTAAATGACAGCTGTCTCTGCTCTCATCC 720
DB 169937 TTGTTTTGCAATGGTCTCTCTATGCTTATGCTTAAATGACAGCTGTCTCTGCTCTCATCC 169996
QY 721 CATGAAGCTCGGTCCAGGCCCTAGGACCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTG 780
DB 169997 CATGAAGCTCGGTCCAGGCCCTAGGACCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTG 170056
```

```
106 781 TCTTATACACAGCCCTCTTCTCTCTTTTATACACACCGCTTTGGCCATCAGTTCAGTC 840
107 170057 TCTTATACACAGCCCTCTTCTCTCTTTTATACACACCGCTTTGGCCATCAGTTCAGTC 170116
108 841 CATATTCAATTTCTTTTGGCCAAATGTTTATCTGCTTTTGGCCACCTGCTCTTAATCTGTG 900
109 170117 CATATTCAATTTCTTTTGGCCAAATGTTTATCTGCTTTTGGCCACCTGCTCTTAATCTGTG 170176
110 901 GTATATGAGGATTAAGACCAACACAGATCGTAAAGAGTTGTTCAGGGTCTTTCAAAAGTCGG 960
111 170177 GTATATGAGGATTAAGACCAACACAGATCGTAAAGAGTTGTTCAGGGTCTTTCAAAAGTCGG 170236
112 961 CAGGAATGGGCATCAAGGCATCTGAG 987
113 170237 CAGGAATGGGCATCAAGGCATCTGAG 170263
```

```
RESULT 3
BD144452
LOCUS BD144452 990 bp DNA linear PAT 17-JAN-2003
DEFINITION Novel G-protein coupled receptors.
ACCESSION BD144452
VERSION BD144452.1 GI:27850210
KEYWORDS JP 2002112793-A/177.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 990)
AUTHORS Haga,T., Takeda,S. and Miyake,N.
TITLE Novel G-protein coupled receptors
JOURNAL Patent: JP 2002112793-A 177 16-APR-2002;
JAPAN SCIENCE AND TECHNOLOGY CORP
COMMENT OS Homo sapiens (human)
PN JP 2002112793-A/177
PF 16-APR-2002
PI 09-FEB-2001 JP 2001034434
PC TATSUYA HAGA, SHIGEKI TAKEDA, NARIKI MIYAKE
PC C12N15/09,A01K67/027,A61K38/00,A61K39/395,A61K39/395,A61K45/00, PC
A61K48/00,
PC A61P43/00,C07K14/705,C07K16/28,C07K19/00,C12N1/15,C12N1/19, PC
C12N1/21,
PC C12N5/10,C12Q1/02,C12Q1/68,G01N33/15,G01N33/50,G01N33/53, PC
G01N33/566//
PC C12P21/08,C12N15/00,A61K37/02,C12N5/00
CC Novel G-protein coupled receptors
FH Key Location/Qualifiers
FT CDS
Location/Qualifiers
1..990
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 99.4%; Score 980.6; DB 6; Length 990;
Best Local Similarity 99.6%; Pred. No. 1.3e-286;
Matches 983; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 ATGTCACGACTTTTGGCCACAACATGGAATCTCTCATCACTGATGTTGACCTTCT 60
DB 1 ATGTCACGACTTTTGGCCACAACATGGAATCTCTCATCACTGATGTTGACCTTCT 60
QY 61 GTCTTCTTCTCTGGGATCCAGGCTGGNACAAATTTTCTTTGGCTCTCACTCCCT 120
DB 61 GTCTTCTTCTCTGGGATCCAGGCTGGNACAAATTTTCTTTGGCTCTCACTCCCT 120
QY 121 GTGTGTGCTTAGGCACAGCACAAATTTGTGGCAATATAACTATTTCTGTTGTGTC 180
DB 121 GTGTGTGCTTAGGCACAGCACAAATTTGTGGCAATATAACTATTTCTGTTGTGTC 180
QY 181 ACTGAACAGCTTTCACAAGCCTGTGATACCTTTTCTGTGTGATGCTCTCAACCATCGAC 240
```


QY 721 CATGAAGCTCGTCCAGGCCCTAGGACCTGTGGTCCCATGCTGTGTCATCTCTCATC 780
DB 722 CATGAAGCTCGTCCAGGCCCTAGGACCTGTGGTCCCATGCTGTGTCATCTCTCATC 781
QY 781 TCTTATACACAGGCCCTCTCTCCCTTTTATACACACCGCTTTGGCCATCAGCTTCCAGTC 840
DB 782 TCTTATACACAGGCCCTCTCTCCCTTTTATACACACCGCTTTGGCCATCAGCTTCCAGTC 841
QY 841 CATATTCACATCTTTTGGCCAAATGTTATCTGCTTTTGGCCACCTGCTCTTAATCCTGTG 900
DB 842 CATATTCACATCTTTTGGCCAAATGTTATCTGCTTTTGGCCACCTGCTCTTAATCCTGTG 901
QY 901 GTATATGAGTTAAGACCAACAGATCCGTAAAGAGTTGTCAGAGTGTTCAAAGTGGG 960
DB 902 GTATATGAGTTAAGACCAACAGATCCGTAAAGAGTTGTCAGAGTGTTCAAAGTGGG 961
QY 961 CAGGGAATGGGCATCAAGGCATCTGAG 987
DB 962 CAGGGAATGGGCATCAAGGCATCTGAG 988

RESULT 5
AX551419 998 bp DNA linear PAT 26-NOV-2002
LOCUS Sequence 37 from Patent WO0250276.
DEFINITION AX551419.1 GI:25814219
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
Li, L., Padigaru, M., Ballinger, R.A., Kekuda, R., Colman, S.D.,
Sciore, P., Smithson, G., Peyman, J.A., Macdougall, J.R., Stone, D.,
Vernet, C.A., Shenoy, S., Gunther, E., Millet, I., Tchernev, V.T.,
Anderson, D., Gusev, V., Malyankar, U.M., Zhong, H., Ellerman, K.E. and
Wolenc, A.
Novel proteins and nucleic acids encoding same
Patent: WO 0250276-A 37 27-JUN-2002;
Curagen Corporation (US)
Location/Qualifiers
1..998
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
2..991
/note="unnamed protein product"
/codon_start=1
/protein_id="CAD57921.1"
/db_xref="GI:25814220"
/translation="MSSTLGHNESPNDVDPSVFFLLGIPGLEQFHLWLSLPVGLG
SATATIGNITIVWATEPVLHPVLFCLMSTIDLAASVSTVPKLLAIFWCGAGHI
GSACIAOFFTHAFOMESTVLLAMAFDPAIACHPLRYATILDTIIAHIGVAHV
RGLMLPCPFLIGLNFQSHVILHTYCEHNAVKLACGDRPNRVGLTRALLVIG
VDFICGLSALSAQAVLRLSHREARKALCTGCGSHVCVILISYTPALFSFTHRFGH
HVPVHIILLANVYLLLPALPNVYGVYTKQIRKRVVRVFSQGGMGKASE"

TITLE
JOURNAL
FEATURES
source
1..998
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
2..991
/note="unnamed protein product"
/codon_start=1
/protein_id="CAD57921.1"
/db_xref="GI:25814220"
/translation="MSSTLGHNESPNDVDPSVFFLLGIPGLEQFHLWLSLPVGLG
SATATIGNITIVWATEPVLHPVLFCLMSTIDLAASVSTVPKLLAIFWCGAGHI
GSACIAOFFTHAFOMESTVLLAMAFDPAIACHPLRYATILDTIIAHIGVAHV
RGLMLPCPFLIGLNFQSHVILHTYCEHNAVKLACGDRPNRVGLTRALLVIG
VDFICGLSALSAQAVLRLSHREARKALCTGCGSHVCVILISYTPALFSFTHRFGH
HVPVHIILLANVYLLLPALPNVYGVYTKQIRKRVVRVFSQGGMGKASE"

CDS
1..998
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
2..991
/note="unnamed protein product"
/codon_start=1
/protein_id="CAD57921.1"
/db_xref="GI:25814220"
/translation="MSSTLGHNESPNDVDPSVFFLLGIPGLEQFHLWLSLPVGLG
SATATIGNITIVWATEPVLHPVLFCLMSTIDLAASVSTVPKLLAIFWCGAGHI
GSACIAOFFTHAFOMESTVLLAMAFDPAIACHPLRYATILDTIIAHIGVAHV
RGLMLPCPFLIGLNFQSHVILHTYCEHNAVKLACGDRPNRVGLTRALLVIG
VDFICGLSALSAQAVLRLSHREARKALCTGCGSHVCVILISYTPALFSFTHRFGH
HVPVHIILLANVYLLLPALPNVYGVYTKQIRKRVVRVFSQGGMGKASE"

ORIGIN
Query Match 99.4%; Score 980.6; DB 6; Length 998;
Best Local Similarity 99.6%; Pred. No. 1.3e-286;
Matches 983; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 1 ATGTCCAGCACTCTTGGCCCAACATGGAATCTCCATCATCAGTGTGACCTTCT 60
DB 2 ATGTCCAGCACTCTTGGCCCAACATGGAATCTCCATCATCAGTGTGACCTTCT 61
QY 61 GTCTTCTCTCTGGGCATCCAGGTCCTGGAACATTTTCATTTGGTCTCTCACCTCCT 120
DB 62 GTCTTCTCTCTGGGCATCCAGGTCCTGGAACATTTTCATTTGGTCTCTCACCTCCT 121

QY 121 GTGTGTGGCTTAGGCACAGCCACAATTTGTGGCAATATAAATAATTTCTGGTTGTTGTC 180
DB 122 GTGTGTGGCTTAGGCACAGCCACAATTTGTGGCAATATAAATAATTTCTGGTTGTTGTC 181
QY 181 ACTGAACAGTCTTGGCAAGCCCTGTGTACCTTTTCTGTGATGCTCTCAACCATCCAC 240
DB 182 ACTGAACAGTCTTGGCAAGCCCTGTGTACCTTTTCTGTGATGCTCTCAACCATCCAC 241
QY 241 TTGGTGGCTCTGTCTCCACAGTTCCCAAGCTATCTGGCTATCTTCTGGTGTGGAGCCGGA 300
DB 242 TTGGTGGCTCTGTCTCCACAGTTCCCAAGCTATCTGGCTATCTTCTGGTGTGGAGCCGGA 301
QY 301 CATATATCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 360
DB 302 CATATATCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 361
QY 361 GAGTCCACTGTGCTACTGGCCATGGCCCTTTGATCGCTACGTGGCCATCTGCCACCATCTC 420
DB 362 GAGTCCACTGTGCTACTGGCCATGGCCCTTTGATCGCTACGTGGCCATCTGCCACCATCTC 421
QY 421 CGCTATGCAACAATCTCTCACTGACACCATCATTTGCCACATAGAGGGGTGGAGCTGTAGTG 480
DB 422 CGCTATGCAACAATCTCTCACTGACACCATCATTTGCCACATAGAGGGGTGGAGCTGTAGTG 481
QY 481 CGAGGCTCCCTGCTCATGCTCCCATGCTCCCTTTTATTTGGGCGTTTGAATCTTCTGCCAA 540
DB 482 CGAGGCTCCCTGCTCATGCTCCCATGCTCCCTTTTATTTGGGCGTTTGAATCTTCTGCCAA 541
QY 541 AGCCATGTATCTTACACAGTACTGTGAGGACATGGCTGTGGTGAAGTGGCCCTGTGGA 600
DB 542 AGCCATGTATCTTACACAGTACTGTGAGGACATGGCTGTGGTGAAGTGGCCCTGTGGA 601
QY 601 GACACAGGCTTAACCGCTGTATGGGCTGACAGCTGACCTGCTGCTGCTGCTGCTGCTGCTG 660
DB 602 GACACAGGCTTAACCGCTGTATGGGCTGACAGCTGACCTGCTGCTGCTGCTGCTGCTGCTG 661
QY 661 TTGTTTTGCATTTGCTCTCTCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
DB 662 TTGTTTTGCATTTGCTCTCTCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 721
QY 721 CATGAAGCTCGTCCAGGCCCTAGGACCTGTGGTTCCTCATGCTGCTGCTGCTGCTGCTGCT 780
DB 722 CATGAAGCTCGTCCAGGCCCTAGGACCTGTGGTTCCTCATGCTGCTGCTGCTGCTGCTGCT 781
QY 781 TCTTATACACAGCCCTCTCTCTCTTTTACACACCGCTTTGGCCATCAGCTTCCAGTC 840
DB 782 TCTTATACACAGCCCTCTCTCTCTTTTACACACCGCTTTGGCCATCAGCTTCCAGTC 841
QY 841 CATATTCACATCTTTTGGCCAAATGTTATCTGCTTTTGGCCACCTGCTCTTAACTCTGTG 900
DB 842 CATATTCACATCTTTTGGCCAAATGTTATCTGCTTTTGGCCACCTGCTCTTAACTCTGTG 901
QY 901 GTATATGAGTTAAGACCAACAGATCCGTAAAGAGTTGTGTCAGGTTGTTTCAAAGTGGG 960
DB 902 GTATATGAGTTAAGACCAACAGATCCGTAAAGAGTTGTGTCAGGTTGTTTCAAAGTGGG 961
QY 961 CAGGGAATGGGCATCAAGGCATCTGAG 987
DB 962 CAGGGAATGGGCATCAAGGCATCTGAG 988

RESULT 6
AX646605
LOCUS Sequence 797 from Patent EPI270724.
DEFINITION AX646605
ACCESSION AX646605.1 GI:28798988
VERSION
KEYWORDS
SOURCE
ORGANISM
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
REFERENCE

Suwa,M., Asai,K., Akiyama,Y. and Aburatani,H.
Guanosine triphosphate-binding protein coupled receptors
Patent: EP 1270724-A. 797 02-JAN-2003;
National Institute of Advanced Industrial Science and Technology
(JP) ; Center for Advanced Science and Technology Incubation, Ltd.
(JP)

Location/Qualifiers
1. .1390
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
201. .1190
/notes="unnamed protein product"
/codon_start=1
/protein_id="CAD69429.1"
/db_xref="GI:28798989"
/db_xref="EMBL:CAD69429"
/translation="MSSTLGHNMESPHNDVDPSVFFLLIGLPGFHLWLSPLVCGGL
GTATTGVNITILVAVTEPVLKPVYFLCLMSTIDLAASVSTPKLAI PWGAGHI
SASACLAQWFFTHAFOMESTVLLAMAFDRYVAICHPLRYATILDTLIIAHIGVAVV
RGSLLMLPCPFLIGRLNFCQSHVILHTYCEHMAVVKLACGDRPNRVYGLTRALLVIG
VDLFCIGLSYLSAQAIVLRLSSHEARSKALGTCGSHVCVILISYTPALFSPFTHRFPH
HVPVHILLANVYLLPPLNPVVVIGVTKQIRKRVVRVVFQSGQGMGIKASE"

ORIGIN
Query Match 99.4%; Score 980.6; DB 6; Length 1390;
Best Local Similarity 99.6%; Pred. No. 1.4e-286;
Matches 983; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 1 ATGTCAGACTCTTGGGCACAACTGAATCTCTCATCACACTGATGTTGACCCCTTCT 60
DB 201 ATGTCAGACTCTTGGGCACAACTGAATCTCTCATCACACTGATGTTGACCCCTTCT 260
QY 61 GTCCTCTTCTCTGGGCATCCAGGCTCGAGCAATTCATTGTCGCTCTCATCTCCCT 120
DB 261 GTCCTCTTCTCTGGGCATCCAGGCTCGAGCAATTCATTGTCGCTCTCATCTCCCT 320
QY 121 GTGTGTGGCTTAGGCACAGCACAATTTGTGGCAATATACTATTCTGTTGTTGTTGCC 180
DB 321 GTGTGTGGCTTAGGCACAGCACAATTTGTGGCAATATACTATTCTGTTGTTGTTGCC 380
QY 181 ACTGAACCAAGCTTTGCAAGCCTGTGTACCTTTTCTGTGATGCTCTCAACCACTGAC 240
DB 381 ACTGAACCAAGCTTTGCAAGCCTGTGTACCTTTTCTGTGATGCTCTCAACCACTGAC 440
QY 241 TTGGCTGCTCTGCTCCACAGTTCCTCAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTG 300
DB 441 TTGGCTGCTCTGCTCCACAGTTCCTCAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTG 500
QY 301 CATATATCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 360
DB 501 CATATATCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 560
QY 361 GAGTCCACTGTGCTAGTGGCCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 420
DB 561 GAGTCCACTGTGCTAGTGGCCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 620
QY 421 CGCTATGCGCAATCTTCACTGACACCATCATTTGCCACATAGGGGTGGGAGCTGTAGT 480
DB 621 CGCTATGCGCAATCTTCACTGACACCATCATTTGCCACATAGGGGTGGGAGCTGTAGT 680
QY 481 CGAGGCTCCCTGCTCATGCTCCCATGCTCCCTTCTTTATTGGGCGTTTGAACCTTGCCAA 540
DB 681 CGAGGCTCCCTGCTCATGCTCCCATGCTCCCTTCTTTATTGGGCGTTTGAACCTTGCCAA 740
QY 541 AGCCATGTGATCTTACACATGCTAGTGGACATGCTGCTGCTGCTGCTGCTGCTGCTG 600
DB 741 AGCCATGTGATCTTACACATGCTAGTGGACATGCTGCTGCTGCTGCTGCTGCTGCTG 800
QY 601 GACACACAGGCTTAACCTGTGTATGGGCTGACAGCTGACAGCTGATGGGTGATGGGGTTGAC 660
DB 801 GACACACAGGCTTAACCTGTGTATGGGCTGACAGCTGACAGCTGATGGGTGATGGGGTTGAC 860

QY 661 TTGTTTTGATGGTCTCTCTCTATGCCCCTAAATGCAAGAGTCTCTCTGCGCTCTCATCC 720
DB 861 TTGTTTTGATGGTCTCTCTCTATGCCCCTAAATGCAAGAGTCTCTCTGCGCTCTCATCC 920
QY 721 CATGAAGCTCGGTCCAAAGGCCCTAGGAGCTGTGGTTCCTCATGCTGTGTGTCATCTCATC 780
DB 921 CATGAAGCTCGGTCCAAAGGCCCTAGGAGCTGTGGTTCCTCATGCTGTGTGTCATCTCATC 980
QY 781 TCTTATACACAGCCCT 840
DB 981 TCTTATACACAGCCCT 1040
QY 841 CATATTACATCTCTTTTGGCCAAATGTTATCTGCTTTTGGCCACCTGCTCTTAATCCTGTG 900
DB 1041 CATATTACATCTCTTTTGGCCAAATGTTATCTGCTTTTGGCCACCTGCTCTTAATCCTGTG 1100
QY 901 GTATATGAGGTTAAAGCAACCAAGATCCGTAAAGAGTTGTCAAGGTGTTTCAAAGTGGG 960
DB 1101 GTATATGAGGTTAAAGCAACCAAGATCCGTAAAGAGTTGTCAAGGTGTTTCAAAGTGGG 1160
QY 961 CAGGGAATGGGCATCAAGGCATCTGAG 987
DB 1161 CAGGGAATGGGCATCAAGGCATCTGAG 1187

RESULT 7
AB065534
LOCUS
DEFINITION
Homo sapiens gene for seven transmembrane helix receptor, complete cds, isolate:CBRC7TM_97.
ACCESSION
AB065534
VERSION
AB065534.1 GI:21928364
KEYWORDS
Homo sapiens (human)
SOURCE
Homo sapiens
ORGANISM
Homo sapiens
REFERENCE
AUTHORS
Suwa,M., Sato,T., Okouchi,I., Arita,M., Futami,K., Matsumoto,S., Tsutsumi,S., Aburatani,H., Asai,K. and Akiyama,Y.
TITLE
Genome-wide discovery and analysis of human seven transmembrane helix receptor genes
JOURNAL
Unpublished
REFERENCE
2 (bases 1 to 1390)
AUTHORS
Suwa,M.
TITLE
Direct Submission
JOURNAL
Submitted (11-JUL-2001) Makiko Suwa, Computational Biology Research Center (CBRC), National Institute of Advanced Industrial Science and Technology (AIST); 2-41-6 Aomi Koto-ku, Tokyo 135-0064, Japan (E-mail: m-suwa@aist.go.jp, URL: http://www.cbrc.jp/)
COMMENT
This sequence is a seven transmembrane helix receptor candidate predicted from the whole human genome sequences using our automated system that contains programs of gene finding (GeneDecoder), sequence search, motif-domain assignment and transmembrane helix prediction.
And the sequence is submitted by the collaborative project between [Computational Biology Research Center (CBRC), National Institute of Advanced Industrial Science and Technology (AIST)] and [Genome Science Division, Research Center for Advanced Science and Technology (RCAST), University of Tokyo].
FEATURES
source
1. .1390
/organism="Homo sapiens"
/mol_type="genomic DNA"
/isolate="CBRC7TM_97"
/db_xref="taxon:9606"
/chromosome="11"
201. .1190
/codon_start=1
/evidence=not_experimental
/product="seven transmembrane helix receptor"
/protein_id="BAC05780.1"

/db_xref="GI:21928365"
/translation="MSSTLGHNMESPNHDTVDPSVFFLLGIPGLEQPHLMSLPCVGL
GTATVGNITLIVVATEPVLKKEVYFLCLMLSTDLAASVSTVPKLLAIFWCCAGHI
SASACLAQMFIFHFCMMESVLLAMAFDRVAICHLRYATILITDIIAHIGVAAY
RGSLLMLPCPELIGLNFQSHVILHTYCEHMAVVKLACGDRNRYVGLTAALLVIG
VDLFCIGLSYALSQAVALRLSSHEARSKALGTGSHVNCVILISYTPALFSPFTRFHG
HVPVHIHILLANVILLPPALNPVYGVKTKQIRKRVVRVFSQSGMGKASE"

ORIGIN

Query Match	99.4%;	Score	980.6;	DB	9;	Length	1390;
Best Local Similarity	99.6%;	Pred.	No. 1.4e-286;				
Matches	983;	Conservative	0;	Mismatches	4;	Indels	0;
Gaps	0;						
Qy	1	ATGTCACAGCACTCTTGGCCACAAATGGAATCTCTCATCACTGATGTTGACCTTCT	60				
Db	201	ATGTCACAGCACTCTTGGCCACAAATGGAATCTCTCATCACTGATGTTGACCTTCT	260				
Qy	61	GTCTTCTTCTCTCTGGGCATCCAGTCTGGAAACAATTTCAATTTGGCTCTCACTCCCT	120				
Db	261	GTCTTCTTCTCTCTGGGCATCCAGGCTCTGGAAACAATTTCAATTTGGTGGCTCTCACTCCCT	320				
Qy	121	GTGTGTGGCTTAGGCACACGCAAAATTTGGGGCAATATAAATTTCTGTGTTGTTGGC	180				
Db	321	GTGTGTGGCTTAGGCACACGCAAAATTTGGGGCAATATAAATTTCTGTGTTGTTGGC	380				
Qy	181	ACTGAACAGCTCTTGCAACAGCCTGTGTACCTTTTCTGTGCATGCTCTCAACCATCGAC	240				
Db	381	ACTGAACAGCTCTTGCAACAGCCTGTGTACCTTTTCTGTGCATGCTCTCAACCATCGAC	440				
Qy	241	TTGGCTGCTCTGTCTCCACAGTTTCCAAAGCTACTGGCTATCTTCTGGTGTGGAGCCGA	300				
Db	441	TTGGCTGCTCTGTCTCCACAGTTTCCAAAGCTACTGGCTATCTTCTGGTGTGGAGCCGA	500				
Qy	301	CATATATCTGCTCTGCTGCTGGCCACATATGTTCTTCATTCATGCTTCTGCAATG	360				
Db	501	CATATATCTGCTCTGCTGCTGGCCACATATGTTCTTCATTCATGCTTCTGCAATG	560				
Qy	361	GAGTCCACTGTGCTACTGGCCATGGCTTTGATCGCTTACGTGGCCATCTGCCACCCACTC	420				
Db	561	GAGTCCACTGTGCTACTGGCCATGGCTTTGATCGCTTACGTGGCCATCTGCCACCCACTC	620				
Qy	421	CGCTATGCCACAACTCTCACTGACACCAATCATTTGCCACATAGGGGTGGCAGTGTAGTG	480				
Db	621	CGCTATGCCACAACTCTCACTGACACCAATCATTTGCCACATAGGGGTGGCAGTGTAGTG	680				
Qy	481	CGAGGCTCCCTGCTCATGCTCCCATGCTCCCTTCTTTATTTGGGCGTTTGAATCTTGCCAA	540				
Db	681	CGAGGCTCCCTGCTCATGCTCCCATGCTCCCTTCTTTATTTGGGCGTTTGAATCTTGCCAA	740				
Qy	541	AGCCATGTGATCTTACACAGCTACTGTGAGCACATGCTGCTGCTGGAAGCTGGCTGTGGA	600				
Db	741	AGCCATGTGATCTTACACAGCTACTGTGAGCACATGCTGCTGGAAGCTGGCTGTGGA	800				
Qy	601	GACACAGGCTTAAACCGTGTGTATGGCTGACAGCTGCACTGTTGGTGTATTTGGGTTGAC	660				
Db	801	GACACAGGCTTAAACCGTGTGTATGGCTGACAGCTGCACTGTTGGTGTATTTGGGTTGAC	860				
Qy	661	TTGTTTGTGATTTGCTCTCTCATGCTTAAATTTGCACAAAGCTGTGCTTGGCTCTCATCC	720				
Db	861	TTGTTTGTGATTTGCTCTCTCATGCTTAAATTTGCACAAAGCTGTGCTTGGCTCTCATCC	920				
Qy	721	CATGAAGCTGGTCTCAAGGCCCTTAGGACCTGTGGTTTCCCATGCTGTGCTCATCCTCATC	780				
Db	921	CATGAAGCTGGTCTCAAGGCCCTTAGGACCTGTGGTTTCCCATGCTGTGCTCATCCTCATC	980				
Qy	781	TCTTATACACAGGCCCTTCTCTCTCTTTTATACACACCGCTTTGGCCATACAGTTCCAGTC	840				
Db	981	TCTTATACACAGGCCCTTCTCTCTCTTTTATACACACCGCTTTGGCCATACAGTTCCAGTC	1040				
Qy	841	CATATTCATATCTTTTGGCCAAATGTTATCTGCTTTTGGCCACTGCTCTTAATCTGTG	900				
Db	1041	CATATTCATATCTTTTGGCCAAATGTTATCTGCTTTTGGCCACTGCTCTTAATCTGTG	1100				

Qy	901	GTATATGAGTTAAGACCAACAGATCGTAAAGAGTTGTCAAGGTGTTTCAAGTGGG	960
Db	1101	GTATATGAGTTAAGACCAACAGATCGTAAAGAGTTGTCAAGGTGTTTCAAGTGGG	1160
Qy	961	CAGGAATGGGCATCAAGGCATCTGAG	987
Db	1161	CAGGAATGGGCATCAAGGCATCTGAG	1187

RESULT 8

AC026090/c
LOCUS
DEFINITION
AC026090
SEQUENCE, 19 unordered pieces.
AC026090
VERSION
HTG; HTGS_PHASE1; HTGS_DRAFT.
KEYWORDS
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 172027)
Waterston,R.H.
The sequence of Homo sapiens clone
Unpublished
2 (bases 1 to 172027)
Waterston,R.H.
Direct Submission
Submitted (19-MAR-2000) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
COMMENT
On Sep 1, 2000 this sequence version replaced gi:9838305.

----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc/index.shtml
----- Project Information -----
Center project name: H.NH0658K18
----- Summary Statistics -----
Sequencing vector: M13; 100%
Sequencing method: plasmid; 0%
Chemistry: Dye-terminator Big Dye; 0% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 160996 bases at least Q40
Consensus quality: 163988 bases at least Q30
Consensus quality: 165626 bases at least Q20
Insert size: 188000; agarose-fp
Insert size: 171387; sum-of-contigs
Quality coverage: 4.33 in Q20 bases; agarose-fp
Quality coverage: 4.82 in Q20 bases; sum-of-contigs

----- NOTE: This is a 'working draft' sequence. It currently
* consists of 19 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

* 1	1526:	contig of 1526 bp in length
* 1527	1626:	gap of unknown length
* 1627	2975:	contig of 1349 bp in length
* 2376	3075:	gap of unknown length
* 3076	5073:	contig of 1998 bp in length
* 5074	5173:	gap of unknown length
* 5174	9317:	contig of 4144 bp in length
* 9318	9417:	gap of unknown length
* 9418	13195:	contig of 3778 bp in length
* 13196	13295:	gap of unknown length
* 13296	16482:	contig of 3187 bp in length
* 16483	16582:	gap of unknown length

*	16583	21965:	contig of 5383 bp in length
*	21966	22065:	gap of unknown length
*	22066	33167:	contig of 11102 bp in length
*	33168	33267:	gap of unknown length
*	33268	43162:	contig of 9895 bp in length
*	43163	43262:	gap of unknown length
*	43263	53157:	contig of 9895 bp in length
*	53158	53257:	gap of unknown length
*	53258	61965:	contig of 8708 bp in length
*	61966	62065:	gap of unknown length
*	62066	70847:	contig of 8782 bp in length
*	70848	70947:	gap of unknown length
*	70948	84297:	contig of 13350 bp in length
*	84298	84397:	gap of unknown length
*	84398	100019:	contig of 15622 bp in length
*	100020	100119:	gap of unknown length
*	100120	114785:	contig of 14666 bp in length
*	114786	114885:	gap of unknown length
*	114886	131872:	contig of 16987 bp in length
*	131873	131973:	gap of unknown length
*	131973	147618:	contig of 15646 bp in length
*	147619	147718:	gap of unknown length
*	147719	171498:	contig of 23780 bp in length
*	171499	171598:	gap of unknown length
*	171599	172027:	contig of 429 bp in length
FEATURES			
source			
1. 172027			
/organism="Homo sapiens"			
/mol_type="genomic DNA"			
/db_xref="taxon:9606"			
/chromosome="11"			
/clone="RP11-658K18"			
1. 1526			
misc_feature	/note="assembly_name:Contig14"		
misc_feature	1627. 2975		
misc_feature	/note="assembly_name:Contig17"		
misc_feature	3076. 5073		
misc_feature	/note="assembly_name:Contig18"		
misc_feature	5174. 9317		
misc_feature	/note="assembly_name:Contig19"		
misc_feature	9418. 13195		
misc_feature	/note="assembly_name:Contig20"		
misc_feature	13296. 16482		
misc_feature	/note="assembly_name:Contig21"		
misc_feature	16583. 21965		
misc_feature	/note="assembly_name:Contig22"		
misc_feature	22066. 33167		
misc_feature	/note="assembly_name:Contig23"		
misc_feature	33268. 43162		
misc_feature	/note="assembly_name:Contig24"		
misc_feature	43263. 53157		
misc_feature	/note="assembly_name:Contig25"		
misc_feature	53258. 61965		
misc_feature	/note="assembly_name:Contig26"		
misc_feature	62066. 70847		
misc_feature	/note="assembly_name:Contig27		
misc_feature	clone_end:17		
misc_feature	vector_side:right		
misc_feature	70948. 84297		
misc_feature	/note="assembly_name:Contig28"		
misc_feature	84398. 100019		
misc_feature	/note="assembly_name:Contig29"		
misc_feature	100120. 114785		
misc_feature	/note="assembly_name:Contig30"		
misc_feature	114886. 131872		
misc_feature	/note="assembly_name:Contig31"		
misc_feature	131973. 147618		
misc_feature	/note="assembly_name:Contig32"		
misc_feature	147719. 171498		
misc_feature	/note="assembly_name:Contig33"		
misc_feature	171599. 172027		
misc_feature	/note="assembly_name:Contig16"		

Cook, A., Cooke, P., DeArellano, K., Dewar, K., Diaz, J.S., Dodge, S.,
Faro, S., Ferreira, P., Fitzhugh, W., Gage, D., Galagan, J., Gardyna, S.,
Ginde, S., Goyette, M., Graham, L., Grand-Pierre, N.,
Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C.,
Kamat, A., Karatas, A., Kells, C., LaRocque, K., Lamazares, R.,
Landers, T., Lehoczyk, J., Levine, R., Lindblad-Toh, K., Liu, G.,
Maclean, C., Macdonald, P., Major, J., Marquis, N., Matthews, C.,
McCarthy, M., McEwan, P., McKernan, K., Meldrum, J., Meneus, L.,
Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R.,
Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neil, D.,
Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V.,
Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P.,
Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schupback, R.,
Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N.,
Strauss, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J.,
Topham, K., Travers, M., Travis, N., Trigilio, J., Vassiliev, H.,
Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G.,
Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Direct Submission
Submitted (10-MAY-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
5 (base 1 to 176597)

AUTHORS
Birren, B., Linton, L., Nusbaum, C., Lander, E., Ali, A., Allen, N.,
Anderson, S., Barna, N., Basien, V., Bloom, T., Boguslavsky, L.,
Boukagaiter, B., Brown, A., Camarata, J., Campopiano, A., Chang, J.,
Chazaro, B., Choepel, Y., Colangelo, M., Collins, S., Collymore, A.,
Cook, A., Cooke, P., DeArellano, K., Dewar, K., Diaz, J.S., Dodge, S.,
Faro, S., Ferreira, P., Fitzhugh, W., Gage, D., Galagan, J., Gardyna, S.,
Ginde, S., Goyette, M., Graham, L., Grand-Pierre, N.,
Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C.,
Kamat, A., Karatas, A., Kells, C., LaRocque, K., Lamazares, R.,
Landers, T., Lehoczyk, J., Levine, R., Lindblad-Toh, K., Liu, G.,
Maclean, C., Macdonald, P., Major, J., Marquis, N., Matthews, C.,
McCarthy, M., McEwan, P., McKernan, K., Meldrum, J., Meneus, L.,
Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R.,
Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neil, D.,
Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V.,
Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P.,
Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schupback, R.,
Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N.,
Strauss, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J.,
Topham, K., Travers, M., Travis, N., Trigilio, J., Vassiliev, H.,
Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G.,
Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Direct Submission
Submitted (13-MAY-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On May 10, 2002 this sequence version replaced gi:2017753.
All repeats were identified using RepeatMasker:
Smit, A.P.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html

----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIER
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu

----- Project Information
Center project name: L26176
Center clone name: 44_D_14

----- Location/Qualifiers
1. 176597
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon,9606"
/chromosome="11"
/map="11"
/clone="RP11-44D14"
/clone_lib="RPC1-11 Human Male BAC"
775..802
/rpt_family="(TTCC)n"
1175..1224
/rpt_family="(CA)n"

repeat_region
repeat_region

FEATURES
source

Disigents (US) ; YEDA RESEARCH AND DEVELOPMENT COMPANY, LTD. (IL)
Location/Qualifiers
1. .963
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="(H38g272 nucleotide)"

FEATURES
source

ORIGIN

Query Match 96.9%; Score 956.6; DB 6; Length 963;
Best Local Similarity 99.6%; Pred. No. 2.7e-279;
Matches 959; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 25 ATGGAATCTCTCATACACTGATGTTGACCCCTTCTGCTCTTCTCTCTGGCATCCCA 84
Db 1 ATGGAATCTCTCATACACTGATGTTGACCCCTTCTGCTCTTCTCTCTGGCATCCCA 60

Qy 85 GGTCTGGAACAATTTCATTTGGCTCTCACTCCCTGTGTGGCTTAGGCACAGCCACA 144
Db 61 GGTCTGGAACAATTTCATTTGGCTCTCACTCCCTGTGTGGCTTAGGCACAGCCACA 120

Qy 145 ATTGCGGCAATATAAATTCATTTGGTGTGTTGCACTGAACCAAGCTTTGCAAGCCT 204
Db 121 ATTGCGGCAATATAAATTCATTTGGTGTGTTGCACTGAACCAAGCTTTGCAAGCCT 180

Qy 205 GTGTACCTTTTCTGTGATGCTCTCAACCATCGACTTGGCTGCTCTCTCCACAGTT 264
Db 181 GTGTACCTTTTCTGTGATGCTCTCAACCATCGACTTGGCTGCTCTCTCCACAGTT 240

Qy 265 CCCAAGCTACTGGCTATCTTCTGGTGTGGAGCGGACATATATCGCTCTGCGCTCG 324
Db 241 CCCAAGCTACTGGCTATCTTCTGGTGTGGAGCGGACATATATCGCTCTGCGCTCG 300

Qy 325 GCACATATGTTCTTCAATTCATGCTTCTGATGAGTGCATGCTGCTACTGCGCATG 384
Db 301 GCACATATGTTCTTCAATTCATGCTTCTGATGAGTGCATGCTGCTACTGCGCATG 360

Qy 385 GCCTTTGATCGCTAGTGGAGCTGGCGCTGTGGAGACACAGCGCTTAACCGTGTAT 444
Db 361 GCCTTTGATCGCTAGTGGAGCTGGCGCTGTGGAGACACAGCGCTTAACCGTGTAT 420

Qy 625 GGGCTGACAGCTGCTGTTGGTCAATTTGGGTTGACTTGTGTTGCAATGGTCTCTCTAT 684
Db 601 GGGCTGACAGCTGCTGTTGGTCAATTTGGGTTGACTTGTGTTGCAATGGTCTCTCTAT 660

Qy 685 GCGCTAAATGCAAGCTGCTCTTGGCTCTCATCCCATGAAAGCTGGTCCAAAGGCCCTA 744
Db 661 GCGCTAAATGCAAGCTGCTCTTGGCTCTCATCCCATGAAAGCTGGTCCAAAGGCCCTA 720

Qy 745 GGGACCTGTGTTCCCATGCTGTGTCATCTCTTATACACAGCGCTCTTCTCC 804
Db 721 GGGACCTGTGTTCCCATGCTGTGTCATCTCTTATACACAGCGCTCTTCTCC 780

Qy 805 TTTTATACACAGCGCTTTGGCCATCAGCTTCCAGTCCATATTCACATTTTGGCCAT 864
Db 781 TTTTATACACAGCGCTTTGGCCATCAGCTTCCAGTCCATATTCACATTTTGGCCAT 840

Qy 865 GTTATCTGTTTGGCACTGCTTAAATCTGTGGTATATGAGTTAAGCCAAACAG 924
Db 841 GTTATCTGTTTGGCACTGCTTAAATCTGTGGTATATGAGTTAAGCCAAACAG 900

Qy 925 ATCCGTAAGAGAGTTGTCAAGGTTCTTCAAGTGGCAGGGAATGGGCATCAAGGCATCT 984
Db 901 ATCCGTAAGAGAGTTGTCAAGGTTCTTCAAGTGGCAGGGAATGGGCATCAAGGCATCT 960

Qy 985 GAG 987
Db 961 GAG 963

RESULT 11
AX318261 966 bp DNA linear PAT 14-DEC-2001
DEFINITION Sequence 42 from Patent WO0190359.
ACCESSION AX318261
VERSION AX318261.1 GI:17900921
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Patterson,C., Tribouley,C.M., Yao,M.G., Griffin,J.A., Thornlon,M.,
Lu,Y., Kallick,D.A., Gandhi,A.R. and Au-Young,J.
TITLE G-protein coupled receptors
JOURNAL Patent: WO 0190359-A 42 29-NOV-2001;
Incyte Genomics, Inc. (US)
FEATURES
Location/Qualifiers
source 1..966
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
/note="Incyte ID No: 7476077CB1"

ORIGIN

Query Match 96.9%; Score 956.6; DB 6; Length 966;
Best Local Similarity 99.6%; Pred. No. 2.7e-279;
Matches 959; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 25 ATGGAATCTCTCATACACTGATGTTGACCCCTTCTGCTCTTCTCTCTGGCATCCCA 84
Db 1 ATGGAATCTCTCATACACTGATGTTGACCCCTTCTGCTCTTCTCTCTGGCATCCCA 60

Qy 85 GGTCTGGAACAATTTCATTTGGCTCTCACTCCCTGTGTGGCTTAGGCACAGCCACA 144
Db 61 GGTCTGGAACAATTTCATTTGGCTCTCACTCCCTGTGTGGCTTAGGCACAGCCACA 120

Qy 145 ATTGCGGCAATATAAATTCATTTGGTGTGTTGCACTGAACCAAGCTTTGCAAGCCT 204
Db 121 ATTGCGGCAATATAAATTCATTTGGTGTGTTGCACTGAACCAAGCTTTGCAAGCCT 180

Qy 205 GTGTACCTTTTCTGTGATGCTCTCAACCATCGACTTGGCTGCTCTCTCCACAGTT 264
Db 181 GTGTACCTTTTCTGTGATGCTCTCAACCATCGACTTGGCTGCTCTCTCCACAGTT 240

Qy 265 CCCAAGCTACTGGCTATCTTCTGGTGTGGAGCGGACATATATCGCTCTGCGCTCG 324
Db 241 CCCAAGCTACTGGCTATCTTCTGGTGTGGAGCGGACATATATCGCTCTGCGCTCG 300

Qy 325 GCACATATGTTCTTCAATTCATGCTTCTGATGAGTGCATGCTGCTACTGCGCATG 384
Db 301 GCACATATGTTCTTCAATTCATGCTTCTGATGAGTGCATGCTGCTACTGCGCATG 360

Qy 385 GCCTTTGATCGCTAGTGGAGCTGGCGCTGTGGAGACACAGCGCTTAACCGTGTAT 444
Db 361 GCCTTTGATCGCTAGTGGAGCTGGCGCTGTGGAGACACAGCGCTTAACCGTGTAT 420

Qy 445 ACCATCATTTGCCCATAGAGGCTGAGCTGTAGTGGAGGCTCCCTGCTCATGCTCCCA 504
Db 421 ACCATCATTTGCCCATAGAGGCTGAGCTGTAGTGGAGGCTCCCTGCTCATGCTCCCA 480

Qy 505 TGTCCCTTCTTATTTGGCGTTTGAACCTTCTGCCAAAGCCATGTCATCTACACAGTAC 564
Db 481 TGTCCCTTCTTATTTGGCGTTTGAACCTTCTGCCAAAGCCATGTCATCTACACAGTAC 540

Qy 565 TGTGAGCACATGCTGTGGTGAAGCTGGCGCTGTGGAGACACAGCGCTTAACCGTGTAT 624
Db 541 TGTGAGCACATGCTGTGGTGAAGCTGGCGCTGTGGAGACACAGCGCTTAACCGTGTAT 600

Qy 625 GGGCTGACAGCTGCTGTTGGTCAATTTGGGTTGACTTGTGTTGCAATGGTCTCTCTAT 684
Db 601 GGGCTGACAGCTGCTGTTGGTCAATTTGGGTTGACTTGTGTTGCAATGGTCTCTCTAT 660

Qy 685 GCGCTAAATGCAAGCTGCTCTTGGCTCTCATCCCATGAAAGCTGGTCCAAAGGCCCTA 744
Db 661 GCGCTAAATGCAAGCTGCTCTTGGCTCTCATCCCATGAAAGCTGGTCCAAAGGCCCTA 720

Qy 745 GGGACCTGTGTTCCCATGCTGTGTCATCTCTTATACACAGCGCTCTTCTCC 804
Db 721 GGGACCTGTGTTCCCATGCTGTGTCATCTCTTATACACAGCGCTCTTCTCC 780

Qy 805 TTTTATACACAGCGCTTTGGCCATCAGCTTCCAGTCCATATTCACATTTTGGCCAT 864
Db 781 TTTTATACACAGCGCTTTGGCCATCAGCTTCCAGTCCATATTCACATTTTGGCCAT 840

Qy 865 GTTATCTGTTTGGCACTGCTTAAATCTGTGGTATATGAGTTAAGCCAAACAG 924
Db 841 GTTATCTGTTTGGCACTGCTTAAATCTGTGGTATATGAGTTAAGCCAAACAG 900

QY 565 TGTGAGCAGATGGCTGTGGTGAAGCTGGCCCTGTGGAGACACAGGCCCTAAACCGTGTGTAT 624
 Db |||||
 QY 541 TGTGAGCAGATGGCTGTGGTGAAGCTGGCCCTGTGGAGACACAGGCCCTAAACCGTGTGTAT 600
 Db |||||
 QY 625 GGGCTGACAGCTGCACTGTGGTCAATGGGGTTGACTTGTGTTGCAATGGTCTCTCCCTAT 684
 Db |||||
 QY 601 GGGCTGACAGCTGCACTGTGGTCAATGGGGTTGACTTGTGTTGCAATGGTCTCTCCCTAT 660
 Db |||||
 QY 685 GCCCTAATTTGACAAAGCTGCTTGGCTCTCTATCCCATGAAGCTCGGTCCAAAGGCCCTTA 744
 Db |||||
 QY 661 GCCCTAAGTGACAAAGCTGCTTGGCTCTCTATCCCATGAAGCTCGGTCCAAAGGCCCTTA 720
 Db |||||
 QY 745 GGGACCTGTGGTTCCCATGTCTGTGTCATCTCTCTATCCCATGAAGCTCGGTCCAAAGGCCCTTA 804
 Db |||||
 QY 721 GGGACCTGTGGTTCCCATGTCTGTGTCATCTCTCTATCCCATGAAGCTCGGTCCAAAGGCCCTTA 780
 Db |||||
 QY 805 TTTTATACACACCGCTTTGGCCATCACGTTCCAGTCCCATATTCACATTTCTTTTGGCCCAAT 864
 Db |||||
 QY 781 TTTTATACACACCGCTTTGGCCATCACGTTCCAGTCCCATATTCACATTTCTTTTGGCCCAAT 840
 Db |||||
 QY 865 GTTATCTGCTTTGGCCACCTGCTTAAATCTGTGGTATATGAGTTAAGACCAAAACAG 924
 Db |||||
 QY 841 GTTATCTGCTTTGGCCACCTGCTTAAATCTGTGGTATATGAGTTAAGACCAAAACAG 900
 Db |||||
 QY 925 ATCCGTAAGAGTTGTCAGGCTGTTTCAAGTGGCAGGGAATGGGCATCAAGGCATCT 984
 Db |||||
 QY 901 ATCCGTAAGAGTTGTCAGGCTGTTTCAAGTGGCAGGGAATGGGCATCAAGGCATCT 960
 Db |||||
 QY 985 GAG 987
 Db |||||
 QY 961 GAG 963
 Db |||||

RESULT 12
 AX448757
 LOCUS
 DEFINITION
 Sequence 423 from Patent WO0224726.
 ACCESSION
 AX448757
 VERSION
 AX448757.1 GI:21697655

KEYWORDS
 SOURCE
 ORGANISM
 Homo sapiens (human)
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 Patent: WO 0224726-A 423 28-MAR-2002;
 ChemCom S.A. (BE)

FEATURES
 Location/Qualifiers

1..966
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"
 1..966
 /note="unnamed protein product"
 /codon_start=1
 /protein_id="CA037694.1"
 /db_xref="GI:21697655"
 /db_xref="REMTREMBL:CA037694"
 /translation="MESPNHTDVPSPVFFLLGIPGLBQPHLWLSLPVCGLTATVGN
 ITILVVATPEVLHPKPVLEFLMLSTIDLASVSTVFKLLAIWFGAGHISASACLAQ
 MFFIHACMESVTLVLAEDRYVAICHPLRYATILITDIIAHIGVAAVVRGSLMLP
 CPELIGRLNFCQSHVILHTCEHMAVVKACGDPTRNRYGLTAALLVIGVDFICGL
 SYALSAQVLRSLSGHSARSLGTGSHVCVILISYTPALFSFTFHRFGHVPVHIH
 LLANVYLLLPALNPVYGVKTKQIRKRVVRVFGSGQGMGKASE"

ORIGIN

Query Match 96.9%; Score 956.6; DB 6; Length 966;
 Best Local Similarity 99.6%; Pred. No. 2.7e-279;
 Matches 959; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

RESULT 13
 AC109341/c
 LOCUS
 DEFINITION

AC109341 202761 bp DNA linear PRI 14-MAY-2002
 Homo sapiens chromosome 11, clone CTD-2504M7, complete sequence.

QY 25 ATGGAATCTCCTCATCAGCTGATGTTGACCCCTTCTGTCTTCTCCTCTGGGCAATCCCA 84
 Db |||||
 QY 1 ATGGAATCTCCTTAATCACTGATGTTGACCCCTTCTGTCTTCTCCTCTGGGCAATCCCA 60
 Db |||||
 QY 85 GGTCTGGAAACAATTTCAATTTGGTCTCTCACTCCCTGTGTGGCTTAGGCACACGCCACA 144
 Db |||||
 QY 61 GGTCTGGAAACAATTTCAATTTGGTCTCTCACTCCCTGTGTGGCTTAGGCACACGCCACA 120
 Db |||||
 QY 145 ATTTGGGCAATATAAATAATTTCTGGTTGTTGGCACTGAACCAAGCTTTGTCACAAGCT 204
 Db |||||
 QY 121 ATTTGGGCAATATAAATAATTTCTGGTTGTTGGCACTGAACCAAGCTTTGTCACAAGCT 180
 Db |||||
 QY 205 GTGTACCTTTTCTGTGCAATGCTCTCAACCAATCGAGCTTGGCTGCTGTCTCCACAGTT 264
 Db |||||
 QY 181 GTGTACCTTTTCTGTGCAATGCTCTCAACCAATCGAGCTTGGCTGCTGTCTCCACAGTT 240
 Db |||||
 QY 265 CCCAAGCTTACTGGCTATCTTCTGGTGTGGAGCCGACATATATCTGCTCTGCCCTGCCCTG 324
 Db |||||
 QY 241 CCCAAGCTTACTGGCTATCTTCTGGTGTGGAGCCGACATATATCTGCTCTGCCCTGCCCTG 300
 Db |||||
 QY 325 GCACATATGTTTCTTCAATTCATGCTTCTGCAATGATGAGTCCACTGTGCTACTGGCCCATG 384
 Db |||||
 QY 301 GCACATATGTTTCTTCAATTCATGCTTCTGCAATGATGAGTCCACTGTGCTACTGGCCCATG 360
 Db |||||
 QY 385 GCCTTTGATCGTACGTGGCCATCTGCCACCACCTCCGCTATGCGCAATCTCTCACTGAC 444
 Db |||||
 QY 361 GCCTTTGATCGTACGTGGCCATCTGCCACCACCTCCGCTATGCGCAATCTCTCACTGAC 420
 Db |||||
 QY 445 ACCATCATTTGCCACATAGGGGTGGCAGCTCTAGTGGGAGGCTCCCTGCTCATGTGCCCA 504
 Db |||||
 QY 421 ACCATCATTTGCCACATAGGGGTGGCAGCTCTAGTGGGAGGCTCCCTGCTCATGTGCCCA 480
 Db |||||
 QY 505 TGTCCCTTCTTTAATTTGGGCGTTTGAACCTTCTGCCAAAGCCAATGTGATCCTACACACGTAC 564
 Db |||||
 QY 481 TGTCCCTTCTTTAATTTGGGCGTTTGAACCTTCTGCCAAAGCCAATGTGATCCTACACACGTAC 540
 Db |||||
 QY 565 TGTGAGCAGATGGCTGTGGTGAAGCTGGCTGTGGAGACACAGGCCCTAAACCGTGTGTAT 624
 Db |||||
 QY 541 TGTGAGCAGATGGCTGTGGTGAAGCTGGCTGTGGAGACACAGGCCCTAAACCGTGTGTAT 600
 Db |||||
 QY 625 GGGCTGACAGCTGCACCTGTTGGTCAATTTGGGTTGACTTGTGTTTGCATTTGGTCTCTCCTAT 684
 Db |||||
 QY 601 GGGCTGACAGCTGCACCTGTTGGTCAATTTGGGTTGACTTGTGTTTGCATTTGGTCTCTCCTAT 660
 Db |||||
 QY 685 GCCCTAATTTGACAAAGCTGCTTCTGCTCTCATGCCATGAAGCTCGGTCCAAAGGCCCTTA 744
 Db |||||
 QY 661 GCCCTAAGTGACAAAGCTGCTTCTGCTCTCATGCCATGAAGCTCGGTCCAAAGGCCCTTA 720
 Db |||||
 QY 745 GGGACCTGTGGTTCCCATGTCTGTGTCATCTCTTATACACAGCCCTCTTCTCC 804
 Db |||||
 QY 721 GGGACCTGTGGTTCCCATGTCTGTGTCATCTCTTATACACAGCCCTCTTCTCC 780
 Db |||||
 QY 805 TTTTATACACACCGCTTTGGCCATCACGTTCCAGTCCCATATTCACATTTCTTTTGGCCCAAT 864
 Db |||||
 QY 781 TTTTATACACACCGCTTTGGCCATCACGTTCCAGTCCCATATTCACATTTCTTTTGGCCCAAT 840
 Db |||||
 QY 865 GTTATCTGCTTTTGGCCACCTGCTCTTAAATCTGTGTGTATATGAGTTAAGACCAAAACAG 924
 Db |||||
 QY 841 GTTATCTGCTTTTGGCCACCTGCTCTTAAATCTGTGTGTATATGAGTTAAGACCAAAACAG 900
 Db |||||
 QY 925 ATCCGTAAGAGTTGTCAGGCTGTTTCAAGTGGCAGGGAATGGGCATCAAGGCATCT 984
 Db |||||
 QY 901 ATCCGTAAGAGTTGTCAGGCTGTTTCAAGTGGCAGGGAATGGGCATCAAGGCATCT 960
 Db |||||
 QY 985 GAG 987
 Db |||||
 QY 961 GAG 963
 Db |||||


```

/rpt_family="HAL1"
complement(13044. .13443)
/rpt_family="L1MA9"
complement(13639. .13939)
/rpt_family="AluSx"
16194. .16222
/rpt_family="AT_rich"
complement(17502. .17712)
/rpt_family="L1MA9"
complement(17688. .18018)
/rpt_family="L1MA9"
complement(18247. .18311)
/rpt_family="L2"
complement(19166. .19280)
/rpt_family="MIR"
19450. .19637
/rpt_family="MIR"
21072. .21116
/rpt_family="L1MB1"
complement(21146. .21534)
/rpt_family="L1MA4"
21536. .21774
/rpt_family="L1MA4"
22101. .22162
/rpt_family="TTCA)n"
23080. .23388
/rpt_family="AluSq"
23817. .23955
/rpt_family="Charlie"
complement(24081. .24150)
/rpt_family="L1MB3"
complement(24151. .24251)
/rpt_family="L1MA1"
24254. .24829
/rpt_family="L1MA1"
complement(24841. .25271)
/rpt_family="L1MB6"
25424. .25503
/rpt_family="(TA)n"
25712. .26331
/rpt_family="L1MA2"
26719. .26754
/rpt_family="AT_rich"
27022. .27136
/rpt_family="(TTATA)n"
28928. .28957
/rpt_family="AT_rich"
29743. .30367
/rpt_family="L1PB3"
complement(30373. .30691)
/rpt_family="L1MB6"
complement(30926. .30987)

```


Maskeri, B., McDowell, J., Mullikin, J.C., Paguirigan, C., Pearson, R., Portnoy, M.E., Prasad, A., Reddix-Dugue, N., Schandler, K., Schueler, M.G., Shah, K., Sison, C., Stantripop, S., Thomas, J.W., Thomas, P.J., Tsipouri, V., Vogt, J.L., Wetherby, K.D., Young, A. and Green, E.D.

NISC Comparative Sequencing Initiative

Unpublished
2 (bases 1 to 239335)

Green, E.D.

Direct Submission

Submitted (11-SEP-2003) NIH Intramural Sequencing Center, 8717

Grovenmont Circle, Gaithersburg, MD 20877, USA

3 (bases 1 to 239335)

Green, E.D.

Direct Submission

Submitted (17-OCT-2003) NIH Intramural Sequencing Center, 8717

Grovenmont Circle, Gaithersburg, MD 20877, USA

On Oct 17, 2003 this sequence version replaced gi:34576313.

----- Genome Center

Center: NIH Intramural Sequencing Center

Center code: NISC

Web site: <http://www.nisc.nih.gov>

Contact: nisc_zoo@nhgri.nih.gov

----- Project Information

Center project name: fgs

Center clone name: 105A13

The sequence data in this record represents an 'enhanced' version of a Phase 2 submission. Specifically, the indicated order and orientation of each sequence contig has been established using one or more of the following: read-pair data from individual subclones, overlaps with neighboring clones, alignment with available reference sequence (e.g., human), and/or confirmation by PCR testing. In addition, the sequence assembly is based on at least 8X average coverage in Q20 bases and has been reviewed to rule out gross misassemblies, the low-quality ends of sequence contigs have been trimmed away, and each base is associated with a Phrap-derived quality score.

----- Summary Statistics

Sequencing vector: plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 238379 bases at least Q40
Consensus quality: 238694 bases at least Q30
Consensus quality: 238929 bases at least Q20
Insert size: 239000; agarose-fp
Insert size: 239035; sum-of-contigs
Quality coverage: 10.78x in Q20 bases; agarose-fp
Quality coverage: 10.78x in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently consists of 4 contigs. Gaps between the contigs are represented as runs of N. The order of the pieces is believed to be correct as given, however the sizes of the gaps between them are based on estimates that have been provided by the submitter.

* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.

* 1 4357: contig of 4357 bp in length
* 4358 4457: gap of unknown length
* 4458 6663: contig of 6206 bp in length
* 6664 66763: gap of unknown length
* 66764 232439: contig of 165676 bp in length
* 232440 232539: gap of unknown length
* 232540 239335: contig of 6796 bp in length.

Location/Qualifiers

1..239335
/organism="Otoleum garnettii"
/mol_type="genomic DNA"
/db_xref="taxon:30611"
/clone="CH256-105A13"

misc_feature
1..4357
/note="assembly_fragment"
clone_end:T7
vector_side:left
4458..6663
/note="assembly_fragment"
66764..232439
/note="assembly_fragment"
232540..239335
/note="assembly_fragment"
clone_end:SP6
vector_side:right

ORIGIN

Query Match 83.3%; Score 822.2; DB 2; Length 239335;
Best Local Similarity 89.6%; Pred. No. 3.9e-238;
Matches 884; Conservative 0; Mismatches 103; Indels 0; Gaps 0;

QY 1 ATGTCCAGCACTCTTGGCCACACATGGAATCTCTCATCACACTGATGTTGACCTTCT 60
Db 9997 ATATCCGGTACTCTTGGCCAGATCATGAAATCTCTTAATCACACACATGACCTTCT 9938
QY 61 GTCTTCTCTCTCTGGGCATCCAGGCTCGAACAAATTTCAITTTGGCTCTCATCCT 120
Db 9937 ATTTCTTCTCTCTTGGGTATTCCAGGCTCTAGAAAAATTTCAITTTGGCTCTCTCATCCT 9878
QY 121 GTGTGTGGCTTTAGGCACAGCCACAAATTTGTGGCAATAAATTTCTCTGTTGTTGTC 180
Db 9877 GTGTGTGGCTTTAGGCACAGCCACAAATTTGTGGCAATAAATTTCTCTGTTGTTGTC 9818
QY 181 ACTGAACAGCTTTGACCAAGCCTGTGTACCTTTTCTGTGCATGCTCTCAACCATCGAC 240
Db 9817 ACTGAGCCAGCCTGTGTACCTTTTCTGTGCATGCTCTCAACCATCGAC 9758
QY 241 TTGCTGCTCTGTCTCCAGAGTCCCAAGTACTGGCTATCTTCTGTGTGGAGCCGGA 300
Db 9757 TTGCTGCTCTGTCTCCAGAGTCCCAAGTACTGGCTATCTTCTGTGTGGAGCTGGA 9698
QY 301 CATATATCTGCTCTGCTGCGCACATATGTTCTTCTTATTTGGCGCTTTGAATCTTGC 360
Db 9697 CATATATCTGCTCTGCTGCGCACAGATGTTCTTCTTATTTGGCGCTTTGAATCTTGC 9638
QY 361 GAGTCCACTGTGTACTGTGGCCATGGCCTTTGATCGCTACGTGGCCATCTGCCACCACTC 420
Db 9637 GAGTCCACTGTGTACTGTGGCCATGGCCTTTGATCGCTACGTGGCCATCTGCCACCACTC 9578
QY 421 CGCTATGCCAATCTCTGACACCATCATTTGCCACATAGGGGTGGAGCTGTAGTG 480
Db 9577 CGCTATGCCAATCTCTGACACCATCATTTGCCACATAGGGGTGGAGCTGTAGTG 9518
QY 481 CGAGGCTCCCTGCTCATGCTCCCATGTCCTTCTTATTTGGCGCTTTGAATCTTGC 540
Db 9517 CGAGGCTCCCTGCTCATGCTCCCATGTCCTTCTTATTTGGCGCTTTGAATCTTGC 9458
QY 541 AGCCATGTGATCTTACACAGCTACTGTGAGCACATGCTGTGGAAGCTGGCTGTGGA 600
Db 9457 AGCCATGTGATCTTACACAGCTACTGTGAGCACATGCTGTGGAAGCTGGCTGTGGA 9398
QY 601 GACACAGGCTTAACCGTGTATGGCTGACAGCTGACATGTTGGTTCATTTGGGGTGGAC 660
Db 9397 GACACAGGCTTAACCGTGTATGGCTGACAGCTGACATGTTGGTTCATTTGGGGTGGAC 9338
QY 661 TTGTTTGGCATTTGCTCTCTTATGCTTATGCTCCCATGTCCTTCTGCTCTCATCC 720
Db 9337 TTGTTTGGCATTTGCTCTCTTATGCTTATGCTCCCATGTCCTTCTGCTCTCATCC 9278
QY 721 CATGAAGCTCGGTCCAAAGGCCCTAGGACCTGTGTTTCCCATGTCCTGTGCTCATC 780
Db 9277 CATGAAGCTCGGTCCAAAGGCCCTAGGACCTGTGTTTCCCATGTCCTGTGCTCATC 9218
QY 781 TCTTATACACAGCCCTCTCTCTCTTTTACACAGCGCTTGGCCATACGTTTCCAGTC 840

Db 9217 TCCTAGACACAGCTCTCTTTTCCTTTTACACATCGCTTTGGCCACACGTTCCACTC 9158
 Qy 841 CATATTCACATTCCTTTTGGCCAAATGTTTATCTGCTTTTGGCCACCTGCTCTTAAATCCCTGTG 900
 Db 9157 CATATTCACATTCCTTTTGGCCAAATGTTTATCTGCTTTCCACCTGCTCTTAAACCTGTGTA 9098
 Qy 901 GTATATCGAGTTAAGACCAACACAGATCCGTAAGAGAGTTGTCAGGGTGTTCAGAGTGGG 960
 Db 9097 GTATATGGGGTTAAGACCAAGAGATTCGTGAAAGGGTTGTCAGAGTGTTCAGAGAGGA 9038
 Qy 961 CAGGAATGGGCATCAAGGCATCTGAG 987
 Db 9037 CAGGAATTGGAATCAAGGCATCTGAG 9011

Search completed: August 27, 2004, 17:37:04
 Job time : 4178 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 27, 2004, 16:10:03 ; Search time 3091 Seconds

(without alignments)
9535.416 Million cell updates/sec

Title: US-10-081-775-1_COPY_537_1523

Perfect score: 987

Sequence: 1 atgtccagcactcttgcca.....tgggcatacaggaatcttgag 987

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: em_estba:*

2: em_esthum:*

3: em_estin:*

4: em_estmu:*

5: em_estov:*

6: em_estpl:*

7: em_estro:*

8: em_estc:*

9: gb_est1:*

10: gb_est2:*

11: gb_est3:*

12: gb_est4:*

13: gb_est5:*

14: gb_est6:*

15: em_estfun:*

16: em_estom:*

17: em_gss_hum:*

18: em_gss_inv:*

19: em_gss_pln:*

20: em_gss_vrt:*

21: em_gss_fun:*

22: em_gss_mam:*

23: em_gss_mus:*

24: em_gss_pro:*

25: em_gss_rod:*

26: em_gss_phg:*

27: em_gss_vrl:*

28: gb_gss1:*

29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	346.6	35.1	948	29	AY401474	AY401474 Mus muscu
2	345	35.0	948	29	AY401472	AY401472 Homo sapi
3	344	34.9	948	29	AY410600	AY410600 Mus muscu
4	340.4	34.5	948	29	AY410598	AY410598 Homo sapi

5	337.2	34.2	948	29	AY410599	AY410599 Pan trogl
6	336.4	34.1	948	29	AY401473	AY401473 Pan trogl
7	331.2	33.6	957	29	AY414225	AY414225 Homo sapi
8	327	33.1	1964	11	AK036356	AK036356 Mus muscu
9	327	33.1	3410	11	AK028467	AK028467 Mus muscu
10	314	31.8	822	29	CC500683	CC500683 CH240_338
11	310	31.4	936	29	AY401478	AY401478 Homo sapi
12	307.8	31.2	954	29	AY414227	AY414227 Mus muscu
13	300.6	30.5	936	29	AY401480	AY401480 Mus muscu
14	297.8	30.2	774	29	CC512357	CC512357 CH240_356
15	294.4	29.8	900	29	AY410595	AY410595 Homo sapi
16	293.6	29.7	966	29	AY405569	AY405569 Mus muscu
17	291.4	29.5	900	29	AY410596	AY410596 Pan trogl
18	284.2	28.8	846	29	AY402272	AY402272 Mus muscu
19	283.6	28.7	966	29	AY405567	AY405567 Homo sapi
20	283	28.7	900	29	AY410597	AY410597 Mus muscu
21	281.6	28.5	846	29	AY402270	AY402270 Homo sapi
22	272.6	27.6	957	29	AY414226	AY414226 Pan trogl
23	270.2	27.4	1086	28	AF101706	AF101706 AF101706
24	256.4	26.0	746	29	AY401479	AY401479 Pan trogl
25	254.6	25.8	758	29	CE341239	CE341239 tigr-gss-
26	251.2	25.5	864	29	CC480759	CC480759 CH240_308
27	238.6	24.2	836	29	CC526526	CC526526 CH240_401
28	236	23.9	798	28	BH082934	BH082934 RPCI-24-9
29	231.4	23.4	939	29	AY402268	AY402268 Pan trogl
30	229.8	23.3	939	29	AY402267	AY402267 Homo sapi
31	229.6	23.3	948	29	AY402752	AY402752 Mus muscu
32	228.8	23.2	820	28	BH046083	BH046083 RPCI-24-2
33	227.8	23.1	966	29	AY401649	AY401649 Homo sapi
34	218.8	22.2	853	29	CC527411	CC527411 CH240_402
35	217.2	22.0	963	29	AY401651	AY401651 Mus muscu
36	216.6	21.9	825	28	AF156720	AF156720 AF156720
37	213.2	21.6	948	29	AY402750	AY402750 Homo sapi
38	213	21.6	701	29	CE070260	CE070260 tigr-gss-
39	211.8	21.5	942	29	AY402281	AY402281 Mus muscu
40	211.4	21.4	709	29	CC503779	CC503779 CH240_343
41	211	21.4	669	29	CC569584	CC569584 CH240_444
42	209.2	21.2	705	28	AF101881	AF101881 AF101881
43	207.8	21.1	786	29	AY414228	AY414228 Homo sapi
44	206.6	20.9	822	29	AY405759	AY405759 Homo sapi
45	203.6	20.6	795	28	BZ256668	BZ256668 CH230-334

ALIGNMENTS

RESULT 1	AY401474	948 bp	DNA	linear	GSS 12-DEC-2003
LOCUS	AY401474	Mus musculus HCM0902 gene, VIRTUAL TRANSCRIPT, partial sequence,			
DEFINITION	AY401474	genomic survey sequence.			
ACCESSION	AY401474				
VERSION	AY401474.1	GI:39757463			
KEYWORDS	GSS.				
SOURCE	Mus musculus (house mouse)				
ORGANISM	Mus musculus				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.				
AUTHORS	1 (bases 1 to 948) Clark, A.G., Gnanowski, S., Nielson, R., Thomas, P., Kejarival, A., Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B., Ferriera, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J., Adams, M.D. and Cargill, M.				
TITLE	Infering nonneutral evolution from human-chimp-mouse orthologous gene trios				
JOURNAL	Science 302 (5652), 1960-1963 (2003)				
PUBMED	14671302				
REFERENCE	2 (bases 1 to 948) Clark, A.G., Gnanowski, S., Nielson, R., Thomas, P., Kejarival, A., Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B., Ferriera, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J., Adams, M.D. and Cargill, M.				
AUTHORS	Direct Submission				
TITLE					

JOURNAL	Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA						
COMMENT	This sequence was made by sequencing genomic exons and ordering them based on alignment.						
FEATURES	Location/Qualifiers						
source	1..948						
gene	/organism="Mus musculus" /mol_type="genomic DNA" /db_xref="taxon:10090" <1..>948 /locus_tag="HMC0902"						
ORIGIN							
Query Match	35.1%;	Score	346.6;	DB 29;	Length	948;	
Best Local Similarity	61.4%;	Pred. No.	3.6e-86;				
Matches	575;	Conservative	0;	Mismatches	359;	Indels 3; Gaps 1;	
Qy	25	ATGGAATCTCTCATCACACTGATGTTCAGCCCTTCTGTTCTTCTCTCCTCGGCAATCCCA	84				
Db	1	ATGCAACACAAACCACGCCATCAGAACCACAAGCTCTTCTTGCTCATGGGATTCCC	60				
Qy	85	GGTCTGGAACAATTTCAATTTGGTCTCTCACTCCCTGTGTGTGGCTTAGGCACACGCCACA	144				
Db	61	GGCCTGGAGGCGTCCCACCTTTTGGATTGCAATTCCTTTCTGTTCCATGTATGCCCCCTAGCA	120				
Qy	145	ATTGTGGCAATAFAAATAATTTCTGGTGTGTTGTGCCACTGAACCAAGTCTTGCACAAGCCT	204				
Db	121	GTTCTTGGCAACATGGCAGTCTCTGTGTGTTGTGGCTCTGAGCCTTCTCATGCAACAACC	180				
Qy	205	GTGTACCTTTTTCTGTGCATGCTCTCAACCATCGACTTGGCTGCTCTGTCTCCACAGTT	264				
Db	181	ATGTACCTGTTCCTGTGCATGCTGTCCAATATTGACCTGTAFACTCTGGACCTTCCACTGTC	240				
Qy	265	CCCAAGCTACTGGGTATCTTCTGTGTGGAGCCGACATATATCTGCCCTCTGCTGCCTG	324				
Db	241	CCCAAGCTCCTCGCACTCTTTTGGCAAATGCAGCTGAGATTGCCTTTTGGAGCCTGTGCT	300				
Qy	325	GCACATATGTTCTTCATTCATGCTTCTTGCAATGATGGAGTCCACTGTGTGTACTGCGCAATG	384				
Db	301	ACCAGATGTTCTTTATCCATGGTTCTCTCGGCTGTAGAATCCGGTATCTCTGCTATCAATG	360				
Qy	385	GCTTTTGATGCTACGTGGCCATCTGCCACCCACCTCCGCTATGCCACAATCTCTACTGCAC	444				
Db	361	GCTTTTGATCCCTACTTGGCCATCTGCCGACCAATTCGACTACGGGTCAATTGCTGTCTCTCA	420				
Qy	445	ACCATCAATPGCCACATAGGGGTGGCAGCTGTAGTGGAGGCTCCCTGCTCATGTCTCCCA	504				
Db	421	GAGTCTGTGACAAAGTTAGGAGCTGCTGCTCTGCTCCGTGGCTCGGGGCTCATGACCCCA	480				
Qy	505	TGTCCTTCTTTATTGGCGGTTTGAACCTTCTGCCAAAGCCATGTGATCCTACACAGGTAC	564				
Db	481	CTTACTCTGCTACTTGGCAAGGCTGAGCTACTG---TGCGCGAGTGGTAGCAATTTCTCTAC	537				
Qy	565	TGTGAGCATGGCTGTGTGAAGCTGGCTCTGGAGACACACAGGCCTTAACCGTGTGTAT	624				
Db	538	TGTGAGCATAGGCTGTGTGAAGCTGGCTCTGGAGGACACACACCAACAAACATATAT	597				
Qy	625	GGGCTGACAGCTGCACTGTGTGTCAATGGGGTTGACTTGTTTTGTCAATGGTCTCTCTAT	684				
Db	598	GGCATCACAGCAGCCACACTGGTGGTAGGAAGTGAATCCATTTGTTATTTGCCATATCTCTAT	657				
Qy	685	GCCCTAATTGCACAAGCTGTCTTCGCTCTCATCCCATGAGCTCGTTCGAAGGCCCTTA	744				
Db	658	GCACCTATCTCTCGGGCTGTGTGTGGGCTCTCTCTCCAAAGAGGCCAGGGCAAGACCTTT	717				
Qy	745	GGGACCTGTGGTTCCTCATGTCTGTGTCTATCTCTCATCTCTTATACACACAGCCCTCTTCTCC	804				
Db	718	GGCATCTGTGGCTCCCACTGGGTGTATCTCTCTCTCTATACACAGGGCTCTTCTCA	777				
Qy	805	TTTTTTACACACCGCTTTTGGCAATACGTTTCAGTCCATATTCATTCATTTCTTTTGGCAAT	864				
Db	778	TTCTATATCCACGCTTTTGGTCAACATGTATACCCAGCAGCTGCACATCTCTCTGGTAC	837				

JOURNAL	Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA									
COMMENT	This sequence was made by sequencing genomic exons and ordering them based on alignment.									
FEATURES	Location/Qualifiers									
source	1..948									
	/organism="Pan troglodytes"									
	/mol_type="genomic DNA"									
gene	<1..>948									
	/db_xref="taxon:9598"									
	/gene="ORS2R1"									
ORIGIN	/locus_tag="HCM3944"									
	Query Match	34.2%	Score	337.2;	DB	29;	Length	948;		
	Best Local Similarity	60.9%;	Pred. No.	1.6e-83;						
	Matches	549;	Conservative	0;	Mismatches	353;	Indels	0;	Gaps	0;
Qy	41	ACACTGATGTGACCCCTTCTGTCTTCTTCCTCTCGCTGGGCATCCAGGTCTCGAACAATTC	101							
Db	20	ACAGCTCTTCTCATCCTGTGCTCTTTCATCTCTGCTTGGAAATCCAGGCCCTGGAGAGTTTCC	79							
Qy	101	ATTTGTGGCTCTCACTCCCTGCTGTGTGGCTTAGGCACAGCCACCAATTCGTGGCAATATAA	160							
Db	80	AGTTTGTGATTGCCCTTCCGTTCTGTGCAATATGCTGTGGCTGTGTGGAATATCA	139							
Qy	161	CTATTCTGGTTGTGTTGCCACTGAACCAAGCTTGTGCACAAGCCTGTGTACCTTTTCTGT	220							
Db	140	CTCTCTCCATGTAATCAGAATTGACCACACCTGCATGAGCCCATGTACCTTCTTCGG	199							
Qy	221	GCATGCTCTCAACCATGACATTGGCTGCTCTGTCTCCACAGTTCCTCAAGTACTGGCTA	280							
Db	200	CCATGCTGGCCATCACTGACCTGGTGGCTCTCCTCTCCACTCAACCTAAGATGTTGGCCA	259							
Qy	281	TCTTCTGGTGTGGAGCCGGAATATATCTGCTCTGCTCTGCTGGCCTGGCGCATATGTTCTTCA	340							
Db	260	TATTCTGGTTTCATGCTCATAGATTTCAGTACCAATGCTGCTGCTCATCCAGGTGTTCTTCA	319							
Qy	341	TTCATGCTCTTCGATGATGAGTGCCACTGTGCTACTGGCCATGGCCCTTGATCGCTACG	400							
Db	320	TCCATGCTTTTCTCTGTGGAGTCTGGGTGCTCATGGCTATGGCCCTGGACTGCTAGC	379							
Qy	401	TGGCCATCTGCCACCCTTCGCTATGCCACAATCCTCACTGACACCATCATFTGCCACA	460							
Db	380	TGGCTATCTGCTTCCCACTCCGACACTTAGCATCCTGACCCCATCGTCTGTGATAAACC	439							
Qy	461	TAGGGTGGCAGCTGTAGTGGAGGCTCCCTGCTCATGCTCCCATGTCCCTTCTTTATG	520							
Db	440	TGGGGACCATCGTGATCTGAGAGGGCTGCTGTGGGTGAGCCCTTCTGCTTCATGGTGT	499							
Qy	521	GGCCTTTGAACCTTCGCCAAAGCCATGTGATCCCTACACACAGTACTGTGAGACATGGCTG	580							
Db	500	CTAGATGCCCCTTCTGCCACACCAAGCCATTCCCACTCATCTGTGAGGCATGGCTG	559							
Qy	581	TGTTGAAGTGGCTGTGGAGACACCAAGCCCTAAACCGTGTGTATGGGCTGACAGTGCAC	640							
Db	560	TGCTGAAGTTGGTGTGCTGATACAAGCATAAAGTCGTGGGTATGGGCTCTTTTGGGCT	619							
Qy	641	TGTTGGTCATTGGGGTTGACCTTGTTTTCGATTGGTCTCTCCTATGCCCTTAATTGCACAAG	700							
Db	620	TCTCTGGCTGGCTTTGATATGATTGTCATTTGGTATGTCATACGTGATGATTTTGAGAG	679							
Qy	701	CTGTCTCTCGGCTCTCATCCCATGAAGCTCGGTCCAGGCCCTTAGGGACCTGTGTTCC	760							
Db	680	CTGTGCTTCAGTTGCCCTCAGGTGAAGCCCGCCTCAAAGCTTTTAGCAACGTCCTCC	739							
Qy	761	ATGTCGTGTGTCATCCTCATCTCTTATACACCAAGCCCTCTCTCTCTTTTATACACCCCT	820							
Db	740	ATATCTGTGTCATCTTGGCTCTTTATATCCAGCCCTTTTCTTCTCTCACTACTGCT	799							
Qy	821	TTGGCCATCAGTTCCAGTCCCATATTCATATCTTTTGGCCAAATGTTTATCTGCTTTTGC	880							
Db	800	TTGGCCATGATGTGCCCGAGTTGTACACATCTCTGTTGCTTAATCTCATCTACTGATAC	859							


```
QY 325 GCACATATGTTCTTCAATCATGCTTCTGCGATGAGTCCACTGTGCTACTGGCCATG 384
Db 301 GCCCAGATGTTCTTTATCCATGGCTTCTCAAGTGTAGAATCTGGTATACTGCTAGCAATG 360
QY 385 GCCTTTGATCGTACGTGGCCATCTGCGACCACTCCGCTATGCGACAAATCCTCACTGAC 444
Db 361 GCCTTGACCGTACTTAGCCATCTGCGGCTCTGCACTATGGTCAATGCTCTCCCA 420
QY 445 ACCATCAATGCCCACAFAGGGGTGGCAGCTGTAGTGGAGGCTCCCTGCTCATGCTCCCA 504
Db 421 GAGTCTGTGNGCAAGCTGGGGGCTGCAGCCNTGCTTCGNGGTTTGGGACTCATGACCCCA 480
QY 505 TGTCCCTCTTTATATGTTGGCGTTTGAACCTTCTGCCAAAGCCATGTATCCTACACAGCTAC 564
Db 481 CTCACCTGCTTACTTGGCAAGACTGNNNTACT---NNNGTCNNNNNGGCCACTCCCTAC 537
QY 565 TGTGAGCACATGCTGTGGTGAAGCTGGCCCTGTGGAGACACAGCCCTAAACCGTGTGTAT 624
Db 538 TNNNAACACATGGCTGTGTGANGCTGGCTTNNNGAGGAACACAGCCAAACACATCTAT 597
QY 625 GGGCTGACAGCTGACATGTTGGTCAATGGGGTTGACTTGTGATTTGCAATGGTCTCTCTAT 684
Db 598 GGCATCAACCGCTGCCACACTGNNNNNGSCACTGACTCCATCTGTATTGCTGTCTCTAT 657
QY 685 GCCCTAATGACAAAGCTGCTCTTCGCTCTCATCCCATGAAGCTCGGTCCCAAGCCCTA 744
Db 658 GCATCATCTCCGGGCTGTGTTAGGCTTCTTCCAAAGGAGGAAGGGCTTAAGACCTTT 717
QY 745 GGGACCTGTGTTTCCCATGTCTGTGTATCTCTCATCTCTTATACACAGCCCTCTTCTCC 804
Db 718 GGCATTTGTGCTCCCACTGGGTGTATCTTCTCTTACACACAGGACTCTTCTCC 777
QY 805 TTTTACACACCGTTTGGCCATCAGCTTCCAGTCCATATTCACATTTCTTTTGGCCAT 864
Db 778 TTCTACACACAGCGTTTGGCAGCAGCTGCGCCGTCACATCCACATCTCTTCTAGCTGAC 837
QY 865 GTTATATCTGTTTGGCCACCTGCTTAATCTCTGTGTATATGAGTTAAGACCAACAG 924
Db 838 CTCTACCTGTTGTGCCACCACTGCTCAACCCCATCATATGATGATGAAGACCAACAG 897
QY 925 ATCGTAAAGAGTTGTGAGGTTGTTTCAAGTGGGC 961
Db 898 ATCTGGATGGGCGCTTCTGAGTGGGCGC 934
```

RESULT 7

```
AY414225
LOCUS
DEFINITION
AY414225
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
```

```
AY414225
Homo sapiens
Homo sapiens (human)
```

```
REFERENCE
AUTHORS
TITLE
```

```
JOURNAL
PUBMED
REFERENCE
AUTHORS
```

```
AY414225
Homo sapiens
Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 957)
Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarial,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
Inferring nonneutral evolution from human-chimp-mouse orthologous
gene trios
Science 302 (5652), 1960-1963 (2003)
14671302
2 (bases 1 to 957)
Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarial,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
Direct Submission
```

```
JOURNAL
COMMENT
FEATURES
source
gene
ORIGIN
```

```
Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
This sequence was made by sequencing genomic exons and ordering
them based on alignment.
Location/Qualifiers
1..957
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
<1..>957
/locus_tag="HCM5153"
```

```
Query Match 33.6%; Score 331.2; DB 29; Length 957;
Best Local Similarity 61.2%; Pred. No. 8e-82;
Matches 553; Conservative 0; Mismatches 348; Indels 3; Gaps 1;
```

```
QY 63 CTTCTTCTCCTCGGCGATCCAGGCTCTGGAACAATTTCAITTTGGTCTCTCACCTCCCTGT 122
Db 42 CTTCTTCTTAACGGGATCCAGGGCTGGAGGCTGCCACTTCTGGATTGCCATCCCTTT 101
QY 123 GTGTGGCTTAGGCACAGCCACAATTTGGGGCAATATTAATTTCTTGGTTGTTGGTCCAC 182
Db 102 CTGTGCCATGTATCTTGTAGCACTGGTTGGAAATGCTGCCCTCATCTGGTCAATGGCAT 161
QY 183 TGAACCAAGTCTTGCAAGCTGTGTACCTTTTCTGTGCATGCTCTCAACCATCGACTT 242
Db 162 GGAACAATGCTCTTATGACACCTATGTACCTTCTCTGCTCTCTCATCTACAGACCT 221
QY 243 GGCTGCCTCTGTCTCCACAGTCTCCAAAGCTACTGGCTATCTTCTGGTGTGGAGCGGACA 302
Db 222 GGCTCTAGTCTACCACTGTGCCCAAGATGCTGCCCATTTTGGTCTCCATGGTGA 281
QY 303 TATATCTGCTCTGCTGCTGCTGGCAGCATATGTTCTTCAATTCATGCTTCTGCATGATGA 362
Db 282 GATTTCCTTTGGTGGATGCTGCGCCAGATGTTTGTGTCCATCTATCTATGCTCTGGA 341
QY 363 GTCCACTGTGCTACTGCGCCATGGCTTTGATGCTAGTGGCCATCTGCCACCCACTCCG 422
Db 342 GTCTCTGATCTACTTCTGCGCATGGCTTTGATAGGTATGTGGCTATCTGTAAACCATTAAG 401
QY 423 CTATGCCACAATCTCTACTGACACCATCATTCGCCACATAGGGGTGGCAGCTGTAGTGC 482
Db 402 GTATACAACCATTTCTCAACCATGCTGTCTAGGAGCAGAAATGGCTTTGTGGCTATTCG 461
QY 483 AGGCTCCCTGCTCATGCTCCCATGCTCCCTTTTATTTGGGGGTTTGAACCTTCTGCCAAG 542
Db 462 TAGTGTGGCTATTGCTCTCCCTTTCTATCTTCTGAGGCGACTCCCTACTGTGGTCA 521
QY 543 CCATGTGATCTCTACACAGCTACTGTGAGCAGCATGGCTGTGTGAAGCTGGCTGTGGAGA 602
Db 522 CCGTGTATGACACACACATCTGTGAGCATATGGGCAATGCCGCACTGGGCTGTGCCAA 581
QY 603 CACCAGGCTTAACCGTGTGTATGGGCTGACAGCTGACACTGTGTGTCATTTGGGGTTGACTT 662
Db 582 CATCACTGTCAATATTGCTATGGCTAACTGTGGCTCTGCTGSCCATGGGACTGGAATTC 641
QY 663 GTTTTGTGATGGTCTCTCTATGCTTAAATTTGACAGAGCTGTCTTGGCTCTCATGCCA 722
Db 642 CATTTCTATTTGCCATTTCTTATGGCTTTATCTCTCATGCACTCTTTCACCTTCCATCTCA 701
QY 723 TGAAGCTCGGTTCGAAGGCTTGGGACTGTGGTTCCTGCTGCTGTGTCATCTCTCATCTC 782
Db 702 TGATGCCAGCACAAGCTCTGAGTACTGTGGGCTCCCAATTTGGCATCATCTCTGGTTT 761
QY 783 TTATACACAGGCTCTTCTCTCTTTTATACACCGCTTTGGCCATCAC---GTTCCAGT 839
Db 762 CTACATCCCTGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 821
QY 840 CCATATTCATCTTTTGGCCATGTTATCTCTCTTTTGGCAGCTGTCTCTTATCTCTCT 899
Db 822 GCATGTGCACATCTTTCTGGCTAATCTATGTGCTGGTCTCTCTCTCTCTCTCTCTCTCT 881
```


Db 668 AGCAATGATCTACTATCTGCTGGCATGCGCTTTGACCGGTACGTGGCTATCTGCCA 727
Qy 414 CCCACTCGGCTATGCCAATCTCTACATGACACACATCATTTGCCACATAGGGGTGGCAGC 473
Db 728 CCCACTCGGCTACGCTGCTGCTCTCAACAATACAGTAACAGTCCAAATAGGATGGTGGC 787
Qy 474 TGTAGTGGAGGCTCCCTGCTCATGCTCCCATGTCCTTCTTTATTTGGCGCTTTGAACCTT 533
Db 788 TCTGGTCCGGGATCCCTATTCCTTTTCCCACTCCCACTACTGATCAAGCAGCATGGCTTT 847
Qy 534 CTGCAAAAGCATGTGATCTCTACACAGCTACTGTGAGCACATGCTGTGGTGAAGCTGGC 593
Db 848 CTGTCACTCCAAATGTCTCTCCCACTCTATTTGTCTCCACAGATGTGATGAAGTGGC 907
Qy 594 CTGTGGAGACACAGGCTCAACCGTGTGTATGGCTGACAGCTGACATGTTGGTCAATGG 653
Db 908 CTATACAGACATATTACCAATGTAGTCTATGGTCTTAAGTCTCACTGCTAGTCAATGGG 967
Qy 654 GGTGACTGTTTGGCATTTGCTCTCTATGCTCCCTAATTCGACAGCTGCTCCGCT 713
Db 968 TGTAGATGTCATGTTCACTCTCTGCTCTCTGCTGCTGATCATAGACGGTTCGCAACT 1027
Qy 714 CTCACTCCCATGAAGCTCGGTCCAAAGCCCTAGGACCTGTGTTCCCATGCTCTGTGTCAT 773
Db 1028 GCCTTCCAAAGTCTGAGCGAGCTAAGGCATTTGGGACTTGTGTGTCACACATTAAGTGGT 1087
Qy 774 CCTCATCTCTTATACACAGCGCTCTTCTCTCTTTTATACACAGCTGTTGGCCATCAGT 833
Db 1088 CTTGGCTCTCTATGTACCACTCAITGGTCTGTGAGTGTGACCGCTTTTGGAAACAGCCT 1147
Qy 834 TCCAGTCCATATTCACATCTTTTGGCCAAAGTTTATCTGCTTTTGGCCACCTGCTCTTAA 893
Db 1148 GGATCCCATTTGTCATGTTCTTAAGGGGAGTCTTACCTGCTGCTGCTCTCTGATGATCA 1207
Qy 894 TCCTGTGTATATGAGTTAAGACCAACAGATCCGTAAGAAAGTGTGACGGGTGTTTCA 953
Db 1208 TCCCATCATCTACGCTGTGTAAGCAACACAGATCAGACACGGTTCCTGCTATGTTCAA 1267
Qy 954 AAGTGGGAGGAAATGGCATCAAGC 980
Db 1268 GATCAGCTGTGACAGGACATTTGAAGC 1294

RESULT 9
AKO28467
LOCUS
DEFINITION Mus musculus 0 day neonate skin cDNA, RIKEN full-length enriched library, clone:4631413D08 product:OLFACTORY RECEPTOR MOR18-2, full insert sequence.
ACCESSION AKO28467
VERSION AKO28467.1 GI:26324423
KEYWORDS HTC; CAP trapper.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE 1
AUTHORS Carninci,P. and Hayashizaki,Y.
TITLE High-efficiency full-length cDNA cloning
JOURNAL Meth. Enzymol. 303, 19-44 (1999)
MEDLINE 99279253
PUBMED 10349636
REFERENCE 2
AUTHORS Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
TITLE Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL Genome Res. 10 (10), 1617-1630 (2000)
MEDLINE 20499374
PUBMED 11042159
REFERENCE 3
AUTHORS Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P.,

Konno,H., Akiyama,J., Nishi,K., Kitsuunai,T., Tashiro,H., Itoh,M., Yumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A., Yanamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K., Fujiwaka,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Matsuoka,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A., and Hayashizaki,Y.
RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multipillar sequencer
Genome Res. 10 (11), 1757-1771 (2000)
20530913
11076861
4
The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.
Functional annotation of a full-length mouse cDNA collection
Nature 409, 685-690 (2001)
5
The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)
6 (bases 1 to 3410)
Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hashizume,W., Hayashida,K., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T., Hori,P., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kasukawa,T., Katoh,H., Kawai,J., Kojima,Y., Kondo,S., Konno,H., Kouda,M., Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M., Nakamura,M., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Ohsato,N., Okazaki,Y., Saito,R., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N., Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Tagami,M., Tagawa,A., Takahashi,F., Takaku-Akahira,S., Takeda,Y., Tanaka,T., Tomaru,A., Toya,T., Yasunishi,A., Muramatsu,M. and Hayashizaki,Y.
Direct Submission
Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.go.jp, URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222, Fax:81-45-503-9216)
cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
Please visit our web site for further details.
URL:http://genome.gsc.riken.go.jp/
URL:http://fantom.gsc.riken.go.jp/.
FEATURES
source
1. 3410
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM_DB:4631413D08"
/db_xref="MGI:23030376"
/db_xref="taxon:10090"
/clone="4631413D08"
/tissue type="skin"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev stages="0 day neonate"
1141..2103
/note="unnamed protein product; OLFACTORY RECEPTOR MOR18-2 (SPR)|AAL60674, evidence: FASTA, 100%ID, 100%length, match=960"
putative"
/codon_start=1
/protein_id="BAC25966.1"
/db_xref="GI:26324424"
/translation="MSSCNFTATFLIGIGLEAHFWFGPFLSMYAVAFNGCIV FIVITRSTSHAPMYFLCMLAAIDLALSTMPKILALFWFDSREITFDACLAQWFF"
CDS


```

QY 830 ACGTTCCAGTCCATATTCACATCTTTTGGCAATGTTTATCTGCTTTTGGCACTGCTC 889
Db |||||
803 AATACACAGGTTACATTCACATCTTGTGTGCAATCTCTATTGATTATCCACCCCTCTC 862
QY |||||
890 TTAATCTCTGTGATATGAGTAAAGACCAACACAGATCCGTAAAGAGTTGTCAAGGTTGT 949
Db |||||
863 TCAACCCCATATTATGAGGTGAGACCAACAGATTTCGAGAGCGAGTGTCTATGTTT 922
QY |||||
950 TT 951
Db ||
923 TT 924

RESULT 12
LOCUS AY414227 954 bp DNA linear GSS 12-DEC-2003
DEFINITION Mus musculus HCM5153 gene, VIRTUAL TRANSCRIPT, partial sequence,
genomic survey sequence.
ACCESSION AY414227
VERSION AY414227.1 GI:39770189
KEYWORDS GSS.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 954)
Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarial,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
Inferring nonneutral evolution from human-chimp-mouse orthologous
gene trios
JOURNAL Science 302 (5652), 1960-1963 (2003)
PUBMED 14671302
REFERENCE 2 (bases 1 to 954)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarial,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE Direct Submission
JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
COMMENT This sequence was made by sequencing genomic exons and ordering
them based on alignment.
FEATURES
source Location/Qualifiers
1..954
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
<1..>954
/locus_tag="HCM5153"

gene
ORIGIN
Query Match 31.2%; Score 307.8; DB 29; Length 954;
Best Local Similarity 59.6%; Pred. No. 3.1e-75;
Matches 538; Conservative 0; Mismatches 362; Indels 3; Gaps 1;

QY 65 TCTTCTCCTGGGATCCAGTCTGGAAACAATTTCATTTGGGCTCTCATCTCCCTGTGT 124
Db |||||
41 TCTTCTCAACTGGGATCCAGGCTGGAGTTGCGCCACTCTGGATTGCCATCCCTTCT 100
QY 125 GTGGCTTAGGCACACCCACCAATTGGGCAATATTAATCTTCTGTTGTTGGCACTG 184
Db |||||
101 GTGTCAATGATGTGGTCTCTGCTTGGGAATGCTGCGCCTCATCTTATCATTTGGGACAG 160
QY 185 AACCACTCTGCACAAAGCTGTGATCTTTTCTGTGCATGCTCTCAACCACTGCACTGG 244
Db |||||
161 AGATGTGCTTACACACCCATGACCTTCTCTCTGCTTCTCTCTCTCTCTCTCTCTCTCT 220
QY 245 CTGCTCTGTCTCCACAGTCTCCCAAGTCTACTGGCTATCTTCTGGTGTGGAGCGGACATA 304
Db |||||
221 CTCTCAGTCCACCACTGTACCTAAAGTTGGCCATCTTATGGCTCCATCTTAATGAGA 280

```

```

QY 305 TATCTGCCTCTGCCTGCTGCGCACATATGTTCTTTCATTCATGCTTCTTGCATGAGT 364
Db |||||
281 TTTCTCTTGGTGGATGCTCGCCAGATGTTTGTGTCACATCTCCATCTATGCTCTTGAGT 340
QY |||||
365 CCAGTGTCTACTGGCCATGCGCTTTGATCGCTACGTGGCCATCTGCCACCCACCTCCGCT 424
Db |||||
341 CCTCGTCTCTTGGCCATGCGCTTTGATCGATATGATAGTATCTGCAACCCACCTGAGAT 400
QY |||||
425 ATGCCAATCTCACTGACACCAATCAATGCCCCAATAGGGGTGGAGCTGTGATGGCAG 484
Db |||||
401 ATACAAACCATCTCAACCATACTGTCTATAGCCCAAAATATCTTGTGGGATAGTCCGTA 460
QY |||||
485 GCTCCCTGCTCATGCTCCCATGCTCCCTTCTTATTTGGGCGTTTGAATCTTGCCTCAAGCC 544
Db |||||
461 GTGGGCTATTGTTCTCCCTTTTCATCTTTTGTGAGCGACTGCCCTACTGTGTGTACC 520
QY |||||
545 ATGTGATCTCTACACACGCTACTGTGAGCACATGCGGTGTGGTGAAGCTGGCCCTGTGAGACA 604
Db |||||
521 GTGTATGATCTCACACATCTGTGAGCATATGGCAATTTGCTAGGCTTGCCTTGGCCAATA 580
QY |||||
605 CCAGGCTAACCGTGTGTATGGGCTGACAGCTGCACTGTGTGGTCAATTTGGGTTGACTTGT 664
Db |||||
581 TTACTGTTAATATATGTTCTATGGGTTGACTGTAGCTTTTGTGCGCATGGGCTGGATTCCA 640
QY |||||
665 TTTGCATTTGGTCTCTCCCTATGCTGAGCAATTTGCACAAGCTGCTTTCGCTCTCATCCCCATG 724
Db |||||
641 TCCTCATTTGCCATCTCCCTATGCTGCTTCCCGGGTGTCTTCCGCTTTCATCCCCGCTG 700
QY |||||
725 AAGCTTCGCTTCAAGGCCCTTAGGGACCTGTGGTTCCTCATGCTGTGTCTCATCTCTCTCT 784
Db |||||
701 ATGCTCAACACAAAGCTCTGAGTACCTGTGGATCTCACTTGGTGTCTATCTGGTTTCT 760
QY |||||
785 ATACACAGCGCTCTCTCTCTCTTTTATACACAGCTTTGG---CCATCAGTTCCAGTCC 841
Db |||||
761 ATATACCTGCTTCTCTCTCTCTTCCCTTCCCTACCCATCGCTTTGGTCACAATCGAGTCCCAAGC 820
QY |||||
842 ATATTCATCTCTTTTGGCCAACTTTATCTGCTTTTGGCCACCTGCTCTTAACTCTGTTGG 901
Db |||||
821 ATGTGCATATCTTCTGCTGAACCTCTATGTGTGGTGGCTTCCCGTGTCTCAATCCCAATCA 880
QY |||||
902 TATATGAGGATTAAAGACAAACAGATCCGTAAGAGTGTTCAGGGTGTTCAGAAAGTGGGC 961
Db |||||
881 TCTAGGTGCAAGAACCAAGAGATTGGAGTCGGCTCTTAAAACTGCTTCACTTGGGGA 940
QY |||||
962 AGG 964
Db |||||
941 AGG 943

RESULT 13
LOCUS AY401480 936 bp DNA linear GSS 12-DEC-2003
DEFINITION Mus musculus HCM0904 gene, VIRTUAL TRANSCRIPT, partial sequence,
genomic survey sequence.
ACCESSION AY401480
VERSION AY401480.1 GI:39757469
KEYWORDS GSS.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 936)
Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarial,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
Inferring nonneutral evolution from human-chimp-mouse orthologous
gene trios
JOURNAL Science 302 (5652), 1960-1963 (2003)
PUBMED 14671302
REFERENCE 2 (bases 1 to 936)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarial,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,

```


TITLE	Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Shinsky,J.J., Adams,M.D. and Cargill,M.									
JOURNAL	Direct Submission									
COMMENT	Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA									
FEATURES	This sequence was made by sequencing genomic exons and ordering them based on alignment.									
	Location/Qualifiers									
	1..936									
source	/organism="Mus musculus"									
gene	/mol_type="genomic DNA"									
	/db_xref="taxon:10090"									
	<1_->936									
ORIGIN	/locus_tag="HCM0904"									
Query Match	30.5%; Score 300.6; DB 29; Length 936;									
Best Local Similarity	59.0%; Pred. No. 3.3e-73;									
Matches	535; Conservative 0; Mismatches 369; Indels 3; Gaps 1;									
Qy	36	TCATCACACTGATGTTGACCCCTTCGTGCTTCTTCTCTCTGGCATCCAGGCTCTGGAACA	95							
Db	12	TAATAGAGCGTTTCTCACCAGATACATTTTCTCATTTGGGATCCAGGCTCTGGAGGA	71							
Qy	96	ATTTCATTTTGGGCTCTCACTCCCTGTGTGTGGCTTAGGCACAGCACAAATTTGGGGCAA	155							
Db	72	AATCCATGCCGTGGATTCTCGCCCTTCTGCTGTATTATCTTTGGGCCCTAATGGTAAA	131							
Qy	156	TATAACTATTCTGGTTGTGTGGCACTGAACCAAGCTCTTGACAAAGCTGTGTACCTTTT	215							
Db	132	TACCATGATCTCGTAGTTATCAAGACAGAGCGTTCTCTCGAGAGCCCATGTTCTACTT	191							
Qy	216	TCGTGATGCTCTCAACCATCGACTTGGCTGCTCTGTCTCCACAGTTTCCCAAGCTACT	275							
Db	192	TCCTGGCCATCCTTTCTTAGTCGACTTGGCTCTTCTTACAACTCTGTGCCCCGTATGCT	251							
Qy	276	GGCTATCTTCTGGTGTGGAGCGGACATATATCTGCCCTCTGCCCTCGCTGGACATATGTT	335							
Db	252	TGGCATATTCTGGTTTGATGCCCATGAATAAACTTGGTGGCTGTGTGGCTCAGATGTT	311							
Qy	336	CTTCATTCAATGCCCTTCTGCATGATGAGAGTCACATGTGCTACTGGCCATGGCTTTGATCG	395							
Db	312	TTTGATCCATGCCCTTTACTGGCATGGAGGCTGAGGTTCTCATGGCCATGGCTTTTGACCG	371							
Qy	396	CTAGCTGGCCATCTGCCACCCACTCCGCTATGCCACAATCTCTACTGACACCATCATTCG	455							
Db	372	TTATGTAGCCATCTGTGCTCCACTCCACTACAAACCATCTTGACATCTCGTGTCTGAT	431							
Qy	456	CCACATAGGGGTGGCAGCTGTAGTGAGAGGCTCCCTGCTCANGCTCCCAATGCCCTTCTT	515							
Db	432	GGGCATCAGCATATGTGTGGTAAATTCGTCCAGCCCTTGTTTATATGCCCAATGGTCTATCT	491							
Qy	516	TATTGGGGCTTTGAACTTCTTGCCAAAGCATGTGATCTCTACACACGTACTCTGAGACAT	575							
Db	492	CATTACCGCTTACCTTTTGTCAAGCTCATGTAAATAGCCCATTCCTATTGTGAACACAT	551							
Qy	576	GGCTGTGGTGAAGCTGGCCCTGTGGAGACACAGAGCCCTAACCGTGTATGGGCTGACAGC	635							
Db	552	GGGCAITGCAAAATGTCTGTGGAGACATCCATATCAACGCTGTATGGCCT--TTT	608							
Qy	636	TGCACTGTGTGTCATTGGGTTGACTTGTGTTTGTGATTGGTCTCTCCTATGCCCTAATTCG	695							
Db	609	TGTGGTCTCTCTCTTTCTTTTGAACCTTGGTACTTATTGGGCATCTCATATGGGTACATCT	668							
Qy	696	ACAAGCTGTCTCTCGCCTCTCATCCCATGAAGCTCGGTCCAAGGCCCTTAGGAGCCTGTGG	755							
Db	669	TCGTGCTGTGTTTTCGCTCCCATCTCAGGATGCACGGTTAAAGCCCTCAGCACATGTGG	728							
Qy	756	TTCCCATGTCTGTGTGATCTCTTATACACAGCCCTCTTCTCCTTTTTCACACA	815							
Db	729	TTCCCATGTGCTGTCTCTGTGTGTTTCTACATCCCATCACTGCTCTCTCTTCTCACTCA	788							
Qy	816	CCGCTTTGGGCATCAGCTTCCAGTCCCATTTACATTTCTTTTGGCCAAATGTTTATCTGCT	875							

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: August 27, 2004, 17:37:09 ; Search time 4663 Seconds
(without alignments)
3058.087 Million cell updates/sec

Title: US-10-081-775-2

Perfect score: 1718

Sequence: 1 MSSTLGHNVSPHHTDVPDPS.....RKRVRVRFQSGQGMKASE 329

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEVEL=xlh
-Q=/cgn2_1/USPTO.spool/US10081775/runat_20082004_170213_11126/app query.fasta_1.519
-DB=genEmbl -QMT=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPEL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10081775@cgn_1_1_3731@runat_20082004_170213_11126 -NCFU=6 -ICPU=3
-NO MAP -LARGQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl:

1: gb_ba:*
2: gb_hg:*
3: gb_in:*
4: gb_lm:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_or:*
21: em_ov:*
22: em_pat:*
23: em_ph:*
24: em_pl:*
25: em_ro:*
26: em_sts:*
27: em_un:*
28: em_un:*

29: em_vi:*
30: em_hg_hum:*
31: em_hg_inv:*
32: em_hg_other:*
33: em_hg_mus:*
34: em_hg_pln:*
35: em_hg_rtd:*
36: em_hg_mam:*
37: em_hg_vrt:*
38: em_sy:*
39: em_hg_hum:*
40: em_hg_mus:*
41: em_hg_others:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1705	99.2	185330	2	AC124272 Homo sapi
2	1705	98.2	203352	2	AC111179 Homo sapi
3	1691	98.4	990	6	BD144452 Novel G-p
4	1691	98.4	998	6	AX551417 Sequence
5	1691	98.4	998	6	AX551419 Sequence
6	1691	98.4	1390	6	AX646605 Sequence
7	1691	98.4	1390	6	AB065534 Homo sapi
8	1691	98.4	172027	2	AC026090 Homo sapi
9	1691	98.4	178597	9	AC116156 Homo sapi
10	1649	96.0	963	6	AX241675 Sequence
11	1649	96.0	966	6	AX318261 Sequence
12	1649	96.0	966	6	AX448757 Sequence
13	1647.5	95.9	202761	9	AC109341 Homo sapi
14	1558	90.7	181386	2	AC147586 Otolenur
15	1558	90.7	239335	2	AC146635 Otolenur
16	1534	89.3	191432	2	AC135109 Mus muscu
17	1534	89.3	191432	2	AC135109 Mus muscu
18	1532	89.2	264026	2	AC113720 Rattus no
19	1517	88.3	966	10	AY317777 Mus muscu
20	1376	80.1	948	6	AX659373 Sequence
21	1224	71.2	781	4	AY355971 Canis fam
22	979.5	57.0	227938	2	AC098008 Rattus no
23	978.5	57.0	1300	6	AX646641 Sequence
24	978.5	57.0	1300	9	AB065856 Homo sapi
25	978.5	57.0	141485	9	AC009758 Homo sapi
26	977.5	56.9	948	10	AY073340 Mus muscu
27	977.5	56.9	948	10	AY317664 Mus muscu
28	977.5	56.9	256243	2	AC102535 Mus muscu
29	962.5	56.0	170338	2	AC138149 Homo sapi
30	951.5	55.4	82761	9	AF391283 Homo sapi
31	950.5	55.3	948	6	AX659367 Sequence
32	936	54.5	951	10	AY317806 Homo sapi
33	936	54.5	236570	2	AC146610 Mus muscu
34	931.5	54.2	942	6	AX241546 Sequence
35	931.5	54.2	985	6	AX241877 Sequence
36	931.5	54.2	990	6	AX448707 Sequence
37	931.5	54.2	990	6	BD144433 Novel G-p
38	931.5	54.2	1013	6	AX702802 Sequence
39	931.5	54.2	1114	6	AX709234 Sequence
40	931.5	54.2	1364	9	AB065819 Homo sapi
41	931.5	54.2	1390	6	AX646627 Sequence
42	931.5	54.2	165729	2	AC009642 Homo sapi
43	930	54.1	235030	2	AC107331 Rattus no
44	927	54.0	954	10	AY073013 Mus muscu
45	927	54.0	954	10	AY317774 Mus muscu

ALIGNMENTS

RESULT 1

AC124272 185330 bp DNA linear HTG 14-SEP-2002
 Locus sapiens chromosome 11 clone RP11-607K3 map 11, WORKING DRAFT
 DEFINITION
 AC124272
 AC124272.2 GI:22857733
 HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 185330)
 Birren,B., Nusbaum,C. and Lander,E.
 Homo sapiens chromosome 11, clone RP11-607K3
 Unpublished
 2 (bases 1 to 185330)
 Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
 Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavskiy,L.,
 Boukhgalter,B., Brown,A., Camarata,J., Campopiano,A., Chang,J.,
 Chazaro,B., Choepel,Y., Colangelo,M., Collins,S., Collymore,A.,
 Cook,A., Cooke,P., DeArelano,K., Dewar,K., Diaz,J.S., Dodge,S.,
 Faro,S., Ferreira,P., Fitzgerald,M., FitzHugh,W., Gage,D.,
 Galagan,J., Gardyna,S., Ginde,S., Gord,S., Goyette,M., Graham,L.,
 Grand-Pierre,N., Hagos,B., Horton,L., Hulme,W., Iliev,I.,
 Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., LaRoque,K.,
 Lamazares,R., Landers,T., Lehotzky,J., Levine,R., Lindblad-Toh,K.,
 Liu,G., MacLean,C., Macdonald,P., Major,J., Marquis,N.,
 Matthews,C., McCarthy,M., McEwan,P., McKernan,K., Melidrim,J.,
 Meneus,L., Mihova,T., Mieng,V., Murphy,T., Naylor,J., Nguyen,C.,
 Nicol,R., Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P.,
 O'Neil,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N.,
 Pollara,V., Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C.,
 Rogov,P., Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S.,
 Schuback,R., Seaman,S., Severi,P., Spencer,B., Stange-Thomann,N.,
 Stojanovic,N., Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S.,
 Theodore,J., Topham,K., Travers,M., Travis,N., Trigilio,J.,
 Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,
 Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
 Direct Submission
 Submitted (14-JUN-2002) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA
 3 (bases 1 to 185330)
 Birren,B., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S.,
 Barna,N., Bastien,V., Bloom,T., Boguslavskiy,L., Boukhgalter,B.,
 Camarata,J., Chang,J., Chazaro,B., Choepel,Y., Collymore,A.,
 Cook,A., Cooke,P., DeArelano,K., Dewar,K., Diaz,J.S., Dodge,S.,
 Faro,S., Ferreira,P., Fitzgerald,M., FitzHugh,W., Gage,D., Galagan,J.,
 Gardyna,S., Gord,S., Graham,L., Grand-Pierre,N., Hagos,B.,
 Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A.,
 Karatas,A., Kells,C., Landers,T., Levine,R., Lindblad-Toh,K.,
 Liu,G., MacLean,C., Macdonald,P., Major,J., Matthews,C.,
 McCarthy,M., Melidrim,J., Meneus,L., Mihova,T., Mieng,V.,
 Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H.,
 O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K.,
 Phunkhang,P., Pierre,N., Raymond,C., Retta,R., Rise,C., Rogov,P.,
 Roman,J., Roy,A., Schauer,S., Schuback,R., Seaman,S., Severi,P.,
 Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Talamas,J.,
 Tesfaye,S., Theodore,J., Topham,K., Travers,M., Vassiliev,H.,
 Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J.,
 Zembek,L., Zimmer,A. and Zody,M.
 Direct Submission
 Submitted (14-SEP-2002) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA
 On Sep 14, 2002 this sequence version replaced gi:21426295.
 All repeats were identified using RepeatMasker:
 Smit, A.F.A. & Green, P. (1996-1997)
 http://ftp.genome.washington.edu/RM/RepeatMasker.html
 ----- Genome Center
 Center: Whitehead Institute/ MIT Center for Genome Research
 Center code: WIBR
 Web site: http://www-seq.wi.mit.edu
 Contact: sequence_submissions@genome.wi.mit.edu
 ----- Project Information

Center project name: L27346
 Center clone name: 607_K_3
 ----- Summary Statistics
 Sequencing vector: Plasmid; n/a; 100% of reads
 Chemistry: Dye-terminator Big Dye; 100% of reads
 Assembly program: Phrap; version 0.960731
 Consensus quality: 184674 bases at least Q40
 Consensus quality: 184934 bases at least Q30
 Consensus quality: 184980 bases at least Q20
 Insert size: 180000; agarose-fp
 Insert size: 185030; sum-of-contents
 Quality coverage: 19.6 in Q20 bases; agarose-fp
 Quality coverage: 19.1 in Q20 bases; sum-of-contents

 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 4 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.
 * 1 120991: contig of 120991 bp in length
 * 120992 121091: gap of 100 bp
 * 121092 132589: contig of 11498 bp in length
 * 132590 132689: gap of 100 bp
 * 132690 153921: contig of 21232 bp in length
 * 153922 154021: gap of 100 bp
 * 154022 185330: contig of 31309 bp in length.
 * 154022 Location/Qualifiers
 * 1 .185330
 * /organism="Homo sapiens"
 * /mol_type="genomic DNA"
 * /db_xref="taxon:9606"
 * /chromosome="11"
 * /map="11"
 * /clone="RP11-607K3"
 * /clone_lib="RPC1-11 Human Male BAC"
 * 1 .120991
 * /note="assembly_fragment"
 * clone_end:SP6
 * vector_side:left
 * 121092 .132589
 * /note="assembly_fragment"
 * 132690 .153921
 * /note="assembly_fragment"
 * 154022 .185330
 * /note="assembly_fragment"
 * misc_feature
 * misc_feature
 * misc_feature
 * misc_feature
 * ORIGIN
 Alignment Scores:
 Pred. No.: 2,68e-150 Length: 185330
 Score: 1705.00 Matches: 327
 Percent Similarity: 99.39% Conservative: 0
 Best Local Similarity: 99.39% Mismatches: 2
 Query Match: 99.24% Indels: 0
 DB: Gaps: 0
 US-10-081-775-2 (1-329) x AC124272 (1-185330)
 QY 1 MetSerThrLeuGlyHisAsnMetGluSerProHisThrAspValaspProSer 20
 Db 3299 ATGTCAGCAGCTTTGGCCACCAACATGGAATCTCTCATCATCATGATGACCTTCT 3358
 QY 21 ValPhePheLeuLeuGlyLeuProGlyLeuGluGlnPheHisLeuTrpLeuSerLeuPro 40
 Db 3359 GTCTTCTCTCTCTGGGCATCCAGGCTCGGAACAATTTCAATTGGGCTCTCACCCT 3418
 QY 41 ValCysGlyLeuGlyThrAlaThrIleValGlyAsnIleThrIleLeuValValAla 60
 Db 3419 GTGTGTGGTTAGGCACACCCACCAATTTGGGCAATATTAATCTTCTGTTGTTGTC 3478
 QY 61 ThrGluProValLeuHisLysProValTyrLeuPheLeuCysMetLeuSerThrIleAsp 80

```

Db      3479  ACTGAACAGTCTTGCACAGCTGTGTACCTTTTCTGTGATGCTCTCAACCATCGAC 3538
QY      81    IeuAlaAaSerValSerThrValProLysLeuLeuAlaIlePheTrpCysGlyAlaGly 100
Db      3539  TTGGCTGCTCTGTCTCCACAGTTCACAGTACTGGCTATCTCTCTGTGTGGAGCCGA 3598
QY      101   HisIleSerAlaSerAlaCysLeuAlaHisMetPhePheIleHisAlaPheCysMetMet 120
Db      3599  CATATATCTGCTCTGCTCCCTGGCCACATATGTTCTTCAATCATGCTCTTGTGCATGATG 3658
QY      121   GluSerThrValLeuLeuAlaMetAlaPheAspArgTyrValAlaIleCysHisProLeu 140
Db      3659  GAGTCACCTGTGCTACTGGCCATGGCTTTGATCGTACGTGGCCATCTGCCACCCACTC 3718
QY      141   ArgTyrAlaThrIleLeuThrAspThrIleAlaHisIleGlyValAlaAlaValVal 160
Db      3719  CGCTATGCCACCAATCTCACTGACACCATCATTTGCCACATAGGGGTGGCAGCTGTAGTG 3778
QY      161   ArgGlySerLeuLeuMetLeuProCysProPhePheIleGlyArgLeuAsnPheCysGln 180
Db      3779  CGAGGCTCCTGCTCATGCTCCATGCTCCCTTCTTTATTTGGCGGTGTGAACCTTCTGCCAA 3838
QY      181   SerHisValIleLeuHisThrTyrCysGluHisMetAlaValValIlysLeuAlaCysGly 200
Db      3839  AGCCATGTGATCTACACAGTACTGTGAGCACATGGCTGTGGTGAAGTGGCTGTGGA 3898
QY      201   AspThrArgProAsnArgValTyrGlyLeuThrAlaAlaLeuLeuValIleGlyValAsp 220
Db      3899  GACACGAGCCCTAAACCGTGTGTATGGCTGACAGCTGCACCTGTGCTATTTGGGTTGAC 3958
QY      221   LeuPheCysIleGlyLeuSerTyrAlaLeuIleAlaGlnAlaValLeuArgLeuSerSer 240
Db      3959  TGTGTTTGATTTGGTCTCTCCATGCTCCCTAAATGACAAAGCTGTCTTGGCTCTCATCC 4018
QY      241   HisGluAlaArgSerIysAlaLeuGlyThrCysGlySerHisValCysValIleLeuIle 260
Db      4019  CATGAAGCTCGTCCCAAGCCCTAGGACCTGTGGTTCCTCAATGTCTGTCTATCTCATC 4078
QY      261   SerTyrThrProAlaLeuPheSerPheThrHisArgPheGlyHisHisValProVal 280
Db      4079  TCTTATACACAGCCCTCTCTCTCTCTTTTACACACCGCTTTGGCCATCACGTTCCAGTC 4138
QY      281   HistIleHisIleLeuLeuAlaAsnValTyrLeuLeuLeuProProAlaLeuAsnProVal 300
Db      4139  CATATTCATCTTTTGGCCAAATGTTATCTGCTTTTGGCCACCTGCTTAAATCCTGTG 4198
QY      301   ValTyrGlyValIysThrIysGlnIleArgIysArgValValArgValPheGlnSerGly 320
Db      4199  GTATATGAGTTAAGACCAACAGATCAGTAAAGAGGTGTACAGGTGTTTCAAGTGGG 4258
QY      321   GlnGlyMetGlyIleLysAlaSerGlu 329
Db      4259  CAGGAATGGGCATCAAGGCATCTGAG 4285

RESULT 2
AC111179
LOCUS      AC111179
DEFINITION Homo sapiens chromosome 11 clone RP11-589G14 map 11, WORKING DRAFT
SEQUENCE, 4 ordered pieces.
ACCESSION AC111179
VERSION   AC111179.2 GI:21699469
KEYWORDS  HTG; HTGS_PHASE2; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE    Homo sapiens (human)
ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 203352)
AUTHORS   Birren,B., Nusbaum,C. and Lander,E.
TITLE     Homo sapiens chromosome 11, clone RP11-589G14
JOURNAL   Unpublished
REFERENCE 2 (bases 1 to 203352)
AUTHORS   Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,

```

```

Anderson,S., Barna,N., Bastien,V., Boguslavkiy,L., Boukhgalter,B.,
Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B.,
Choepe,Y., Colangelo,M., Collins,S., Collamore,A., Cook,A.,
Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S.,
Ferrelta,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S.,
Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,
Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
Kamat,A., Karatas,A., Kells,C., LaRocque,K., Lamazares,R.,
Landers,T., Lehoczy,J., Levine,R., Liu,G., MacLean,C.,
Macdonald,P., Major,J., Matkuis,N., Matthews,C., McCarthy,M.,
McEwan,P., McKernan,K., Meldrim,J., Meneus,L., Mihova,T.,
Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C.,
Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J.,
Peterson,K., Phunkhang,P., Pierre,N., Pollara,V., Raymond,C.,
Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P., Roman,J.,
Rosett,M., Roy,A., Santos,R., Schauer,S., Schupback,R., Seaman,S.,
Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
Topham,K., Travers,M., Travis,N., Triggilio,J., Vassiliev,H.,
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
Direct Submission
Submitted (18-FEB-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 203352)
Birren,B., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S.,
Barna,N., Bastien,V., Bloom,T., Boguslavkiy,L., Boukhgalter,B.,
Camarata,J., Chang,J., Chazaro,B., Choepe,Y., Collamore,A.,
Cooke,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
Faro,S., Ferreira,P., FitzGerald,M., Gage,D., Galagan,J.,
Gardyna,S., Gord,S., Graham,L., Grand-Pierre,N., Hagos,B.,
Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A.,
Karatas,A., Kells,C., Landers,T., Levine,R., Lindblad-Toh,K.,
Liu,G., MacLean,C., Macdonald,P., Major,J., Matthews,C.,
McCarthy,M., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V.,
Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H.,
O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K.,
Phunkhang,P., Pierre,N., Raymond,C., Retta,R., Rise,C., Rogov,P.,
Roman,J., Roy,A., Schauer,S., Schupback,R., Seaman,S., Severy,P.,
Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Talamas,J.,
Tesfaye,S., Theodore,J., Topham,K., Travers,M., Vassiliev,H.,
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J.,
Zembek,L., Zimmer,A. and Zody,M.
Direct Submission
Submitted (06-AUG-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Jul 5, 2002 this sequence version replaced gi:18699944.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L25440
Center clone name: 589 G.14
----- Summary Statistics
Sequencing vector: Plasmid; n/a; 100% of reads
Assembly: Dye-terminator Big Dye; 100% of reads
Program: Phrap; version 0.960731
Consensus quality: 201894 bases at least Q40
Consensus quality: 202347 bases at least Q30
Consensus quality: 202672 bases at least Q20
Insert size: 200000; agarose-fp
Insert size: 203052; sum-of-contigs
Quality coverage: 14.6 in Q20 bases; agarose-fp
Quality coverage: 14.4 in Q20 bases; sum-of-contigs
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 4 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces

```

TITLE
JOURNALREFERENCE
AUTHORSTITLE
JOURNAL
COMMENT

* is believed to be correct as given, however the sizes
 * of the gaps between them are based on estimates that have
 * provided by the submitter.

* This sequence will be replaced

* by the finished sequence as soon as it is available and

* the accession number will be preserved.

* 1 70426: contig of 70426 bp in length

* 70427 70526: gap of 100 bp

* 70527 78730: contig of 8204 bp in length

* 78731 78930: gap of 100 bp

* 78931 137440: contig of 58610 bp in length

* 137441 137540: gap of 100 bp

* 137541 203352: contig of 65812 bp in length.

FEATURES

source

1..203352

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

/chromosome="11"

/map="11"

/clone="RP11-589G14"

/clone_lib="RPC1-11 Human Male BAC"

1..70426

/note="assembly_fragment"

clone_end:SP6

vector_side:left

70527..78730

/note="assembly_fragment"

78831..137440

/note="assembly_fragment"

137541..203352

/note="assembly_fragment"

clone_end:P7

vector_side:right

ORIGIN

Alignment Scores:
 Pred. No.: 2,998-150 Length: 203352
 Score: 1705.00 Matches: 327
 Percent Similarity: 99.39% Conservative: 0
 Best Local Similarity: 99.39% Mismatches: 2
 Query Match: 99.24% Indels: 0
 DB: 2 Gaps: 0

US-10-081-775-2 (1-329) x AC111179 (1-203352)

QY 1 MetSerThrLeuGlyHisAsnMetGluSerProHisHisThrAspValAspProSer 20
 DB 169277 ATGTCCAGCACTCTTTGGCCACCAACATGGAATCTCTCATCACACTGATGTTGACCTTCT 169336

QY 21 ValPheLeuLeuGlyLeuProGlyLeuGluGlnPheHisLeuTrpLeuSerLeuPro 40
 DB 169337 GTCCTTCTCTCTGGGATCCAGGCTGGGAACAAATTCATTGTGGCTCTCACTCCCT 169396

QY 41 ValCysGlyLeuGlyThrAlaThrIleValGlyAsnIleThrIleLeuValValAla 60
 DB 169397 GTGTGTGGCTTAGGCACACCAATTTGGGCAATATAACTATTCTGTGTGTGGC 169456

QY 61 ThrGluProValLeuHisIysProValTyLeuPheLeuCysMetLeuSerThrIleAsp 80
 DB 169457 ACTGAACCACTTGGCACAAGCTGTGTACCTTTTCTGTGATGCTCTCAACCACTCGAC 169516

QY 81 LeuAlaAlaSerValSerThrValProIysLeuLeuAlaIlePheTrpCysGlyAlaGly 100
 DB 169517 TTGGCTGGCTCTGTCTCCACAGTTCACAGCTACTGGCTATCTTCTGATGATG 169576

QY 101 HisIleSerAlaSerAlaCysLeuAlaHisMetPhePheIleHisAlaPheCysMetMet 120
 DB 169577 CATATATCTGCTCTGCTGCTGGCAGCATATGTTCTTCATTCATGCTCTGATGATG 169636

QY 121 GluSerThrValLeuLeuAlaMetAlaPheAspArgTyValAlaIleCysHisProLeu 140
 DB 169637 GAGTCCACTGTGCTACTGGCCATGGCCTTTTGATCGCTAGTGGCCATCTGCCACCCTC 169696

QY 141 ArgTyAlaThrIleLeuThrAspThrIleAlaHisIleGlyValAlaValVal 160
 DB 169697 CGCTATGCCCAATCTCTCACTGACACCAATATTGCCCAATAGGGGTGGAGCTGTAGTG 169756

QY 161 ArgGlySerLeuLeuMetLeuProCysProPhePheIleGlyArgLeuAsnPheCysGln 180
 DB 169757 CGAGGCTCCCTGCTCATGCTCCCATGTCCTTTTATTGGCGCTTTGAACCTTCTGCCAA 169816

QY 181 SerHisValIleLeuHisThrTyTyCysGluHisMetAlaValValValLeuAlaCysGly 200
 DB 169817 AGCCATGTGATCTTACACACGCTACTGTGAGCACATGGCTGTGGTGAAGCTGGCCTGTGGA 169876

QY 201 AspThrArgProAsnArgValTyGlyLeuThrAlaAlaLeuLeuValIleGlyValAsp 220
 DB 169877 GACACCAGGCTTACCCGTTGTATGGGCTGACAGCTGACCTGTGGTGTATGGGCTTGAC 169936

QY 221 LeuPheCysIleGlyLeuSerTyAlaLeuLeuIleAlaGlnAlaValLeuArgIleuSerSer 240
 DB 169937 TTGTTTTGCAATTGGTCTCTCTATGCTTAAATTGCAAGCTGTCTCTTGGCTCTCATCC 169996

QY 241 HisGluAlaArgSerIysAlaLeuGlyThrCysGlySerHisValCysValIleLeuIle 260
 DB 169997 CATGAAGCTCGTCCAGGCTTAGGACCTGTGGTTCCTCATGTGTGTCTCATCTCATC 170056

QY 261 SerTyThrProAlaLeuPheSerPheThrHisArgPheGlyHisHisValProVal 280
 DB 170057 TCTTATACACAGCCCTCTCTCTTTTATACACCCGCTTTGGCCTCATCGTTCAGTC 170116

QY 281 HisIleHisIleLeuLeuAlaAsnValTyLeuLeuLeuProProAlaLeuAsnProVal 300
 DB 170117 CATATTCACTTCTTTTGGCCAAATGTTTATCTCTCTTTTGGCCACCTGCTTAACTGTG 170176

QY 301 ValTyGlyValIysThrIysGlnIleArgIysArgValValArgValPheGlnSerGly 320
 DB 170177 GTATATGAGTTAAGACCAACACATCAGTAAAGAGGTGTGAGGTGTTTCAAGTGGG 170236

QY 321 GlnGlyMetGlyIleLysAlaSerGlu 329

DB 170237 CAGGGAATGGGATCAGGCATCTGAG 170263

RESULT 3
 BD144452
 LOCUS BD144452 990 bp DNA linear PAT 17-JAN-2003
 DEFINITION Novel G-protein coupled receptors.
 ACCESSION BD144452
 VERSION BD144452.1 GI:27850210
 KEYWORDS JP 2002112793-A/177.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 990)
 AUTHORS Haga, T., Takeda, S. and Miyake, N.
 TITLE Novel G-protein coupled receptors
 JOURNAL Patent: JP 2002112793-A 177 16-APR-2002;
 JAPAN SCIENCE AND TECHNOLOGY CORP
 COMMENT OS Homo sapiens (human)
 FN JP 2002112793-A/177
 PD 16-APR-2002
 PI 03-FEB-2001 JP 2001034434
 PI TATSUYA HAGA, SHIGEKI TAKEDA, NARIKI MIYAKE
 PC
 C12N15/09, A01K67/027, A61K38/00, A61K39/395, A61K39/395, A61K45/00, PC
 A61K48/00,
 PC A61P43/00, C07K14/705, C07K16/28, C07K19/00, C12N1/15, C12N1/19, PC
 C12N1/21,
 PC C12N5/10, C12Q1/02, C12Q1/68, G01N33/15, G01N33/50, G01N33/53, PC
 G01N33/566//
 PC C12P21/08, C12N15/00, A61K37/02, C12N5/00
 CC Novel G-protein coupled receptors
 FH Key Location/Qualifiers
 FT CDS (1)..(990).

```
FEATURES
  source
    Location/Qualifiers
      1..990
        /organism="Homo sapiens"
        /mol_type="genomic DNA"
        /db_xref="taxon:9606"

ORIGIN
Alignment Scores:
  Pred. No.: 1,31e-151 Length: 990
  Score: 1691.00 Matches: 325
  Percent Similarity: 99.09% Conservative: 1
  Best Local Similarity: 98.78% Mismatches: 3
  Query Match: 98.43% Indels: 0
  DB: 6 Gaps: 0

US-10-081-775-2 (1-329) x BD144452 (1-990)

QY 1 MetSerSerThrLeuGlyHisAsnMetGluSerProHisHisThrAspValAspProSer 20
DB 1 ATGTCAGACACTCTTGGCCACACATGGAATCTCCTAATCACACTGATGTTGACCCCTTCT 60
QY 21 ValPhePheLeuGlyLeuProGlyLeuGluGlnPheHisLeuTrpLeuSerLeuPro 40
DB 61 GTCTTCCTCCTCGGCGATCCAGGCTCGGACCAATTTCAATTTGTGGCTCTCACTCCCT 120
QY 41 ValCysGlyLeuGlyThrAlaThrIleValGlyAsnIleThrIleLeuValValAla 60
DB 121 GGTGTGGCTTAGGCACACCAATTTGGGCAATATTAATCTTCTGTTGTGTGTGCC 180
QY 61 ThrGluProValLeuHisLysProValTyrLeuPheLeuPheCysMetLeuSerThrIleAsp 80
DB 181 ACTGAACCAAGCTCTGCACAAAGCTGTGTACCTTTTCTGTGATGCTCTCAACCATCGAC 240
QY 81 LeuAlaAlaSerValSerThrValProLysLeuLeuAlaIlePheTrpCysGlyAlaGly 100
DB 241 TTGGCTGCTCTGTCTCCACAGTTCACCAAGTACTGGCTATCTTCTGGTGTGGAGCCGA 300
QY 101 HisIleSerAlaSerAlaCysLeuAlaHisMetPhePheIleHisAlaPheCysMetMet 120
DB 301 CATATATCTGCTCTGCTGCTGCTGCGACAGATGTTCTTCAATTCATGCTCTTGTGATGATG 360
QY 121 GluSerThrValLeuLeuAlaMetAlaPheAspArgTyrValAlaIleCysHisProLeu 140
DB 361 GAGTCCACTGTCTACTGGCCATGCTTGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 420
QY 141 ArgTyrAlaThrIleLeuThrAspThrIleIleAlaHisIleGlyValAlaAlaValVal 160
DB 421 CCTATGCCCAATCCTCACTGACACCATCATTTGCCCAATAGTGGGCTGCGAGCTGTAGTG 480
QY 161 ArgGlySerLeuLeuMetLeuProCysProPhePheIleGlyArgLeuAsnPheCysGln 180
DB 481 CGAGGCTCCCTGCTCATGCTCCATGCTCCCTTCTTATTTGGGCGTTGAATCTTCCGCAA 540
QY 181 SerHisValIleLeuHisThrTyrCysGluHisMetAlaValValLysLeuAlaCysGly 200
DB 541 AGCCATGTGATCCTATACACAGTACTGTGAGCAGATGCTGCTGCTGCTGCTGCTGCTGCTG 600
QY 201 AspThrArgProAsnArgValTyrGlyLeuThrAlaAlaLeuLeuValIleGlyValAsp 220
DB 601 GACACAGGCTTAACCGGTGTATGGGTGACAGTGCATGCTGCTGCTGCTGCTGCTGCTGCTG 660
QY 221 LeuPheCysIleGlyLeuSerTyrAlaLeuIleAlaGlnAlaValLeuArgLeuSerSer 240
DB 661 TTGTTTTCATTTGGTCTCTCCTATGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 720
QY 241 HisGluAlaArgSerLysAlaLeuGlyThrCysGlySerHisValCysValIleLeuIle 260
DB 721 CATGAAGCTCGGTCACAGGCTTAGGACCTGTGGTTCCTTCTTCTTCTTCTTCTTCTTCTTCT 780
QY 261 SerTyrThrProAlaLeuPheSerPhePheThrHisArgPheGlyHisValProVal 280
DB 781 TCTTATACACGAGCCCTTCTCTCTTTTATACACAGCCCTTTGGCCATCAGCTTCCAGTC 840

FEATURES
  source
    Location/Qualifiers
      1..990
        /organism="Homo sapiens"
        /mol_type="genomic DNA"
        /db_xref="taxon:9606"

ORIGIN
Alignment Scores:
  Pred. No.: 1.32e-151 Length: 998
  Score: 1691.00 Matches: 325
  Percent Similarity: 99.09% Conservative: 1
  Best Local Similarity: 98.78% Mismatches: 3
  Query Match: 98.43% Indels: 0
  DB: 6 Gaps: 0

US-10-081-775-2 (1-329) x AX551417 (1-998)

QY 1 MetSerSerThrLeuGlyHisAsnMetGluSerProHisHisThrAspValAspProSer 20
DB 2 ATGTCCAGCACTCTTGGCCACACATGGAATCTCCTAATCACACTGATGTTGACCCCTTCT 61
QY 21 ValPhePheLeuGlyIleProGlyLeuGluGlnPheHisLeuTrpLeuSerLeuPro 40
DB 62 GTCTTCCTCCTCGGCGATCCAGGCTCGGACCAATTTCAATTTGTGGCTCTCACTCCCT 121
QY 41 ValCysGlyLeuGlyThrAlaThrIleValGlyAsnIleThrIleLeuValValAla 60
DB 122 GTGTGTGGCTTAGGCACACCAATTTGGGCAATATTAATCTTCTGTTGTGTGTGCC 181
```



```
QY 61 ThrGluProValLeuHisIysProValTyrIleuPheLeuCysMetLeuSerThrIleAsp 80
DB 182 ACTGAACAGCTGTGCACAAAGCTGTACCTTTTCTGTGCATGCTCTCAACCATCGAC 241

QY 81 LeuAlaAlaSerValSerThrValProLysLeuLeuAlaIlePheTyrCysGlyAlaGly 100
DB 242 TTGGCTGCTGTCTCCACAGTTCACAGCTACTGGCTACTCTTCTGGTGTGAGCCGGA 301

QY 101 HisIleSerAlaSerAlaCysLeuAlaHisMetPhePheIleHisAlaPheCysMetMet 120
DB 302 CATATATCTGCCTCTGCCTGCTGCACAGATGTTCTTCAATCATGCTCTTGCATGATG 361

QY 121 GluSerThrValLeuAlaMetAlaPheAspArgTyrValAlaIleCysHisProLeu 140
DB 362 GAGTCCACTGTGCTACTGGCCATGGCTTTGATCGCTACGTGGCCATCTGCCACCACATC 421

QY 141 ArgTyrAlaThrIleLeuThrAspThrIleAlaHisIleGlyValAlaAlaValVal 160
DB 422 CGCTATGCCAATCTCTCACTGACACCATCAATGGCCACATAGGGGTGGCAGCTGTAGTG 481

QY 161 ArgGlySerLeuLeuMetLeuProCysProPhePheIleGlyArgIleuAsnPheCysGln 180
DB 482 CGAGGCTCCCTGCTCATGCTCCATGCTCCCTTCTTATTTGGGCGTTTGAACCTTCCGCAA 541

QY 181 SerHisValIleLeuHisThrTyrCysGluHisMetAlaValValIleCysGly 200
DB 542 AGCCATGTGATCTCTACACAGTACTGTGAGCATGCTGGTGTGAGGTGGCCCTGTGGA 601

QY 201 AspThrArgProAsnArgValTyrGlyLeuThrAlaAlaLeuLeuValIleGlyValAsp 220
DB 602 GACACAGGCTTAACCGTGTGATGGCTGACAGCTGCTGTGTCTATTTGGGTTCAC 661

QY 221 LeuPheCysIleGlyLeuSerTyrAlaLeuIleAlaGlnAlaValLeuArgLeuSerSer 240
DB 662 TTGTTTTGCAATGGTCTCTCTATGCTTATGCTTAAAGTGCACAGCTGTCTCTGCTCATCC 721

QY 241 HisGluAlaArgSerIysAlaLeuGlyThrCysGlySerHisValCysValIleLeuIle 260
DB 722 CATGAAGCTCGGTCCAGGCCCTTAGGACCTGTGGTTCCATGTCGTGTCTATCTCATC 781

QY 261 SerTyrThrProAlaLeuPheSerPhePheThrHisArgPheGlyHisValProVal 280
DB 782 TCTATACACAGCCCTCTCTCTCTTTTACACAGCTTTGGCCATCACCTTCAGTC 841

QY 281 HisIleHisIleLeuAlaLeuValTyrLeuLeuLeuProProAlaLeuAsnProVal 300
DB 842 CATATTACATCTTTTGGCCAAATGTTATCTGCTTTTGCCACCTGCTCTTAATCTGTG 901

QY 301 ValTyrGlyValLysThrLysGlnIleArgLysArgValValArgValPheGlnSerGly 320
DB 902 GTATATGGAGTTAAGACCAACAGATCCGTAAAGAGTTGTACAGGGTGTTTCAAGTGGG 961

QY 321 GlnGlyMetGlyIleLysAlaSerGlu 329
DB 962 CAGGGAATGGGCATCAAGGCATCTGAG 988
```

```
RESULT 5
AX551419
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
```

```
Sequence 37 from Patent WO0250276.
AX551419.1 GI:25814219
Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
```

```
Li, L., Padigar, M., Ballinger, R.A., Kekuda, R., Colman, S.D.,
Sciore, P., Smithson, G., Peyton, J.A., Macdougall, J.R., Stone, D.,
Vernet, C.A., Shenoy, S., Gunther, E., Millet, I., Tchernev, V.T.,
Anderson, D., Gusev, V., Malyankar, U.M., Zhong, H., Ellerman, K.E. and
Wolenc, A.
```

```
TITLE Novel proteins and nucleic acids encoding same
JOURNAL Patent: WO 0250276-A 37 27-JUN-2002;
Curagen Corporation (US)
FEATURES
Location/Qualifiers
source
1..998
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
CDS
2..991
/note="unnamed protein product"
/codon_start=1
/protein_id="CAD57921.1"
/db_xref="GI:25814220"
/db_xref="REMBL:CAD57921"
/translation="MSSLTGHMSPNHTDVPVFFLLGIPGLBQPHLMLSLPYCGGL
GTATVGNITILVVVATEPVLHKPYVLELMLSTIDLAASVSTVPKLLAIWCGAGHI
SASCLAQMFPIHAFCMESVLLAMAFDVAICHPRYATILDTLIDTIIAHIGAAVVG
RGLMLPCLPFLIQRLNQCQSHVILHTYCEHMAVVKLACGDTNRNVGLTAAALVIG
VDLFCIGLSYALSAQVLRLSHEARSALAGTCSHVCVILISYTPALFSFTFRFGH
HVPVHIHLLANVYLLPPALNPVYVVKTKQIRKRVVRVFSQGGMGIKASE"
```

ORIGIN

```
Alignment Scores:
Pred. No.: 1.32e-151 Length: 998
Score: 1691.00 Matches: 325
Percent Similarity: 99.0% Conservative: 1
Best Local Similarity: 98.7% Mismatches: 3
Query Match: 98.4% Indels: 0
DB: 6 Gaps: 0
```

US-10-081-775-2 (1-329) x AX551419 (1-998)

```
QY 1 MetSerSerThrLeuGlyHisAsnMetGluSerProHisHisThrAspValAspProSer 20
DB 2 ATGTCACAGCTCTTGGCCACCAACATGGAATCTCTAATACACATGATGTGACCTTCT 61

QY 21 ValPhePheLeuLeuGlyIleProGlyLeuGluGlnPheHisLeuTrpLeuSerLeuPro 40
DB 62 GTCTTCTTCTCTGGSCATCCAGGTCTGGAAACAATTCATTTTGGTCTCTCACTCCCT 121

QY 41 ValCysGlyLeuGlyThrAlaThrIleValGlyAsnIleThrIleLeuValValAla 60
DB 122 GTGTGTGGCTTAGGCACAGCCACAATTTGGGCAATATAAATATTCTGTGTGTGTGTC 181

QY 61 ThrGluProValLeuHisIysProValTyrLeuPheLeuCysMetLeuSerThrIleAsp 80
DB 182 ACTGAACAGCTGTGCACAGCTGTGTACCTTTTCTGTGATGCTCTCAACCATCGAC 241

QY 81 LeuAlaAlaSerValSerThrValProLysLeuLeuAlaIlePheTyrCysGlyAlaGly 100
DB 242 TTGGCTGCTGTCTCTCCACAGTTCACAGCTTCCCAAGCTACTGGCTATCTTCTGTGTGAGCCGGA 301

QY 101 HisIleSerAlaSerAlaCysLeuAlaHisMetPhePheIleHisAlaPheCysMetMet 120
DB 302 CATATATCTGCCTCTGCCTGCTGCACAGATGTTCTTCAATCATGCTCTTGCATGATG 361

QY 121 GluSerThrValLeuLeuAlaMetAlaPheAspArgTyrValAlaIleCysHisProLeu 140
DB 362 GAGTCCACTGTGCTACTGGCCATGGCTTTGATCGCTACGTGGCCATCTGCCACCACATC 421

QY 141 ArgTyrAlaThrIleLeuThrAspThrIleAlaHisIleGlyValAlaAlaValVal 160
DB 422 CGCTATGCCAATCTCTCACTGACACCATCAATGGCCACATAGGGGTGGCAGCTGTAGTG 481

QY 161 ArgGlySerLeuLeuMetLeuProCysProPhePheIleGlyArgIleuAsnPheCysGln 180
DB 482 CGAGGCTCCCTGCTCATGCTCCATGCTCCCTTCTTATTTGGGCGTTTGAACCTTCCGCAA 541

QY 181 SerHisValIleLeuHisThrTyrCysGluHisMetAlaValValIleCysGly 200
DB 542 AGCCATGTGATCTCTACACAGTACTGTGAGCATGCTGGTGTGAGGTGGCCCTGTGGA 601

QY 201 AspThrArgProAsnArgValTyrGlyLeuThrAlaAlaLeuLeuValIleGlyValAsp 220
```

```

Db      602  GACACAGGCCAAACCGTGTGTATGGCTGACAGCTGCATGTTGCTCATTTGGGGTTGAC 661
Qy      221  LeuPheCysIleGlyLeuSerTyrAlaLeuIleAlaGlnAlaValLeuArgLeuSerSer 240
Db      662  TTGTTTTCATTTGCTCTCTCTCTATGCGCTAAAGTGCACAAGCTGCTCTCGCCTCTCATCC 721
Qy      241  HisGluAlaArgSerLysAlaLeuGlyThrCysGlySerHisValCysValIleLeuIle 260
Db      722  CATGAAGCTCGGTCCAGGCCCTAGGGACCTGTGGTTCCTCATGCTGTGTCTCTCATC 781
Qy      261  SerTyrThrProAlaLeuPheSerPheThrHisArgPheGlyHisValProVal 280
Db      782  TCTTATACACAGCCCTCTCTCTCTTTTACACACCGCTTTGGCCATCACGTTCCAGTC 841
Qy      281  HisIleHisIleLeuLeuAlaValTyrLeuLeuLeuProProAlaLeuAsnProVal 300
Db      842  CATATTTCATCTTTTGGCCCAATGTTATCTGCTTTTGGCCACCTGCTCTTAATCTCTG 901
Qy      301  ValTyrGlyValIleLysThrLysGlnIleArgLysArgValValArgValPheGlnSerGly 320
Db      902  GTATATGGAGTTAAGACCAACAGATCCCTTAAGAGTGTGTGAGGTTGTTTCAAAGTGGG 961
Qy      321  GlnGlyMetGlyIleLysAlaSerGlu 329
Db      962  CAGGGAATGGGCATCAAGGCATCTGAG 988

RESULT 6
AX646605
LOCUS      AX646605      1390 bp      DNA      linear      PAT 04-MAR-2003
DEFINITION      Sequence 797 from Patent EP1270724.
ACCESSION      AX646605
VERSION      AX646605.1 GI:28798988
KEYWORDS
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
REFERENCE
AUTHORS      Suwa, M., Asai, K., Akiyama, Y. and Aburatani, H.
TITLE      Guanosine triphosphate-binding protein coupled receptors
JOURNAL      Patent: EP 1270724-A 797 02-JAN-2003;
National Institute of Advanced Industrial Science and Technology
(JP) ; Center for Advanced Science and Technology Incubation, Ltd.
(JP)

FEATURES
source      Location/Qualifiers
1..1390
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
201..1190
/note="unnamed protein product"
/codon_start=1
/protein_id="CAD69429.1"
/db_xref="GI:28798989"
/db_xref="REMBL:CAD69429"
/translation="MSSTLGHMSPHSDVDPSVFFLLGIPGLEQLHMLSLPVCGL
GTATIGNVITLVVAFVPEVLHPKVIYFLCMLSTIDLAASVSTPKLLAIFWCGAGHI
SASACLAQMFIHAFWMESTVLLAMAFDHYVAICHLRYATILDTIHAIGVAAV
RGSLLMLPCPLIGRLNFCQSHVILHTYCEHMAVVKLACGDRPNRVGLTAALLVIG
VDLFCIGLSVALSAQAVRLSSHEARSKALGTCGSHVCVLIISYTPALFSPFTHFGH
HVPVHIILLANVYLLPPALNPVYGVKTKQIRKVRVFPQSGQGMKASE"

ORIGIN
Alignment Scores:
Pred. No.:      1.94e-151      Length:      1390
Score:      1691.00      Matches:      325
Percent Similarity:      99.09%      Conservative:      1
Best Local Similarity:      98.78%      Mismatches:      3
Query Match:      98.43%      Indels:      0
DB:      6      Gaps:      0

US-10-081-775-2 (1-329) x AX646605 (1-1390)

```

RESULT 7
AB065534
LOCUS
DEFINITION

AB065534 1390 bp DNA linear PRI 23-JUL-2002
Homo sapiens gene for seven transmembrane helix receptor, complete

Qy 1 MetSerSerThrIleuGlyHisAsnMetGluSerProHisHisThrAspValAspProSer 20
Db 201 ATGTCAGCACCTTTGGCCACAACATGGAATCTCTTAATCACACTGATGTTGACCTTCT 260
Qy 21 ValPhePheLeuLeuGlyIleProGlyLeuGluGlnPheHisLeuTrpLeuSerLeuPro 40
Db 261 GTCCTTCTCTCTCTGGGCATCCAGGCTCTGGAACAATTTCAATTTGGGCTCTCACTCCCT 320
Qy 41 ValCysGlyLeuGlyThrAlaThrIleValGlyAsnIleThrIleLeuValValAla 60
Db 321 GTGTGTGGCTTAGGCACAGCCCAATTTGGGCAATATTAATCTATTCTGGTGTGTGTGCC 380
Qy 61 ThrGluProValLeuHisLysProValTyrLeuPheLeuCysMetLeuSerThrIleAsp 80
Db 381 ACTGAACCAAGTCTTGCACAAGCCTGTGTACCTTTTCTGTGCATGCTCTCAACCACTCGAC 440
Qy 81 LeuAlaAlaSerValSerThrValProLysLeuLeuAlaIlePheTrpCysGlyAlaGly 100
Db 441 TTGGCTGCTCTGTCTCCACAGTTCACAGTCTCCCAAGTACTGGCTATCTCTGTGTGTGAGCCGA 500
Qy 101 HisIleSerAlaSerAlaCysLeuAlaHisMetPhePheIleHisAlaPheCysMetMet 120
Db 501 CATATATCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 560
Qy 121 GluSerThrValLeuLeuAlaMetAlaPheAspArgTyrValAlaIleCysHisProLeu 140
Db 561 GAGTCCACTGTGCTACTGGCCATGGCTTTGATCGCTAGCTGGCCATCTGCCACCCACTC 620
Qy 141 ArgTyrAlaThrIleLeuThrAspThrIleIleAlaHisIleGlyValAlaAlaValVal 160
Db 621 CGCTATGCCCAATCTCTCACTGACACCATCATATGCCACATAGCCGCTGAGCTGTAGTG 680
Qy 161 ArgGlySerLeuLeuMetLeuProCysProPhePheIleGlyArgLeuAsnPheCysGln 180
Db 681 CGAGCTCTCTGCTCATGCTCCATGCTCCCTTCTTATTTGGCGCTTGAACCTTCTGCCAA 740
Qy 181 SerHisValIleLeuHisThrTyrCysGluHisMetAlaValValLysLeuAlaCysGly 200
Db 741 AGCCATGTGATCCTACACAGTACTGTGAGCACATGCTGTGGTGAAGCTGGCCTGTGGA 800
Qy 201 AspThrArgProAsnArgValTyrGlyLeuThrAlaAlaLeuLeuValIleGlyValAsp 220
Db 801 GACACACAGGCTAAACCGTGTATGGGCTGACAGTGCACCTGTTGGTCAITGGGGTTGAC 860
Qy 221 LeuPheCysIleGlyLeuSerTyrAlaLeuIleAlaGlnAlaValLeuArgLeuSerSer 240
Db 861 TTGTTTTCATTTGCTCTCTCTATGCCCCCTAAGTGCACAGCTGCTCTTGGCTCTCTCATCC 920
Qy 241 HisGluAlaArgSerLysAlaLeuGlyThrCysGlySerHisValCysValIleLeuIle 260
Db 921 CATGAAGCTCGGTCCAGGCCCTAGGGACCTGTGTTCTCCCATGCTGTGTCTCATCTCATC 980
Qy 261 SerTyrThrProAlaLeuPheSerPhePheThrHisArgPheGlyHisHisValProVal 280
Db 981 TCTTATACACAGCCCTCTCTCTCTTTTACACACCGCTTTGGCCATCAGCTTCCAGTC 1040
Qy 281 HisIleHisIleLeuLeuAlaAsnValTyrLeuLeuLeuProProAlaLeuAsnProVal 300
Db 1041 CATATTTCATCTTTTGGCCAAATGTTAATCTGCTTTTGGCCACCTGCTCTTAATCTCTGTG 1100
Qy 301 ValTyrGlyValLysThrLysGlnIleArgLysArgValValArgValPheGlnSerGly 320
Db 1101 GTATATGGAGTTAAGACCAACAGATCCGTAAGAGAGTGTGTCAGGGGTGTTTCAAAGTGGG 1160
Qy 321 GlnGlyMetGlyIleLysAlaSerGlu 329
Db 1161 CAGGGAATGGGCATCAAGGCATCTGAG 1187

cds, isolate:CBRC7TM_97.	Db	321	GTGTGTGGCTTAGGCACAGCCACAATTTGTGGCAATATAACTATTCTGGTGTGTGTGGCC	380
AB065534	QY	61	ThrGluProValLeuHisLysProValTyrLeuPheLeuCysMetLeuSerThrIleAsp	80
AB065534.1 GI:21928364	Db	381	ACTGAACAGTCTTGGCAACAGCTGTGTACCTTTTCTGTGCATGCTCTCAACCATCGAC	440
Homo sapiens (human)	QY	81	LeuAlaAlaSerValSerThrValProLysLeuLeuAlaIlePheTrpCysGlyAlaGly	100
Homo sapiens	Db	441	TTGGCTGCTCTGTCTCCACAGTTCCTCAAGCTACTGGCTATCTTCTGTGTGGAGCCGA	500
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	QY	101	HisLeuSerAlaSerAlaCysLeuAlaHisMetPhePheIleHisAlaPheCysMetMet	120
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	Db	501	CATATATCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	560
Genome-wide discovery and analysis of human seven transmembrane	QY	121	GluSerThrValLeuLeuAlaMetAlaPheAspArgTyrValAlaIleCysHisProLeu	140
helix receptor genes	Db	561	GAGTCCATGTGTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	620
Unpublished	QY	141	ArgTyrAlaThrIleLeuThrAspThrIleAlaHisIleGlyValAlaAlaValVal	160
Suwa, M.	Db	621	CGCTATGCCACAATCTCTCACTGACACCATCATTTGCCACATAGGGGTGGCAGCTGT	680
Direct Submission	QY	161	ArgGlySerLeuLeuMetLeuProCysProPhePheIleGlyArgLeuAsnPheCysGln	180
Submitted (11-JUL-2001) Makiko Suwa, Computational Biology Research	Db	681	CGAGGCTCCCTGCTCATGCTCCCATGTCCTTCTTATTTGGGCGTTTGAACCTTCTG	740
Center (CBRC), National Institute of Advanced Industrial Science	QY	181	SerHisValIleLeuHisThrTyrCysGluHisMetAlaValValLysLeuAlaCysGly	200
and Technology (AIST); 2-41-6 Aomi Koto-ku, Tokyo 135-0064, Japan	Db	741	AGCCATGTGATCCTACACAGTACTGTGAGCAGCATGGTGTGGTGAAGCTGGCCTGTGA	800
(E-mail:m-suwa@aist.go.jp, URL:http://www.cbrc.jp/,	QY	201	AspThrArgProAsnArgValTyrGlyLeuThrAlaAlaLeuLeuValIleGlyValAsp	220
Tel:81-3-3599-8080, Fax:81-3-3599-8081)	Db	801	GACACCGGCTTAACCGTGTGTATGGGCTGACAGCTGCATGTTGGTTCATTTGGGGTTG	860
This sequence is a seven transmembrane helix receptor candidate	QY	221	LeuPheCysIleGlyLeuSerTyrAlaLeuIleAlaGlnAlaValLeuAsnSerSer	240
predicted from the whole human genome sequences using our automated	Db	861	TTGTTTTGCAATGGTCTCTCTCTATGCCCTTAAGTGACAAAGCTGTCTTGGCCTCTCAT	920
system that contains programs of gene	QY	241	HisGluAlaAlaArgSerLysAlaLeuGlyThrCysGlySerHisValCysValIleLeu	260
finding (GeneDecoder), sequence search, motif-domain assignment and	Db	921	CATGAAGCTCGTCCAGGCCCTAGGACCTGTGTGTTCCATGTCGTGTCTCTCTCTCT	980
transmembrane helix prediction.	QY	261	SerTyrThrProAlaLeuPheSerPheThrHisArgPheGlyHisHisValProVal	280
And the sequence is submitted by the collaborative project between	Db	981	TCTTATACACAGCCCTCTCTCTCTTTTACACACCGCTTTGGCCATCAGCTTCCAGTC	1040
[Computational Biology Research Center (CBRC), National Institute	QY	281	HisIleHisIleLeuLeuAlaAsnValTyrLeuLeuLeuProProAlaLeuAsnProVal	300
of Advanced Industrial Science and Technology (AIST)] and [Genome	Db	1041	CATATTCACATCTTTTGGCCCAATGTTTATCTGCTTTTGGCCACCTGCTCTTAACTCTG	1100
Science Division, Research Center for Advanced Science and	QY	301	ValTyrGlyValLysThrLysGlnIleArgLysArgValValArgValPheGlnSerGly	320
Technology (RCST), University of Tokyo].	Db	1101	GTATATGGAGTTAAGACCAACAGATCCGTAAAGAGTTGTTCAGGGGTGTTTCAAAGTGG	1160
Location/Qualifiers	QY	321	GinglyMetGlyIleLysAlaSerGlu	329
1. .1390	Db	1161	CAGGGAATGGGCATCAAGGCACTGTAG	1187
/organism="Homo sapiens"	RESULT 8			
/mol_type="genomic DNA"	AC026090/c			
/isolate="CBRC7TM_97"	LOCUS			
/db_xref="taxon:9606"	DEFINITION			
/chromosome="11"	Homo sapiens chromosome 11 clone RP11-658K18, WORKING DRAFT			
201. .1190	SEQUENCE, 19 unordered pieces.			
/codon_start=1	AC026090			
/evidence=not experimental	172027 bp DNA linear HTG 01-SEP-2000			
/product="seven transmembrane helix receptor"	AC026090			
/protein_id="BAC05780.1"	Accession			
/db_xref="GI:21928365"	VERSION			
/translation="MSSTLGHNESPNTDVPVFFLLGIFGLSQFHLSLPLVPCGL	KEYWORDS			
GTATVGNITIVVATEPVLHPVFLCLMLSTDLAASVSTVPKLAIFWCGAGHI	HTG; HTGS_PHASE1; HTGS_DRAFT.			
SASACLAQWFFTHAFOMESTVLLAMAPDYVAICPLRYATILDTIIAHIGVAIV	Homo sapiens (human)			
PGSLMLPCPLTIGRLNFCQSHVILHTYCEHMAVKLAGDTRPNRVGLTAAALVIG	Homo sapiens			
VDLFCIGLSYALSAQAVLSSHEARSKALCTCGSHVCVILISYTPALFSEFTFRFGH	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
HVPVHIHILLANVILLPPALNPVTVGVTKQIRKRVVRVFGSQGMGKASE"	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
	1 (bases 1 to 172027)			
	Waterston, R.H.			
	The sequence of Homo sapiens clone			
FEATURES				
source				
1. .1390				
/organism="Homo sapiens"				
/mol_type="genomic DNA"				
/isolate="CBRC7TM_97"				
/db_xref="taxon:9606"				
/chromosome="11"				
201. .1190				
/codon_start=1				
/evidence=not experimental				
/product="seven transmembrane helix receptor"				
/protein_id="BAC05780.1"				
/db_xref="GI:21928365"				
/translation="MSSTLGHNESPNTDVPVFFLLGIFGLSQFHLSLPLVPCGL				
GTATVGNITIVVATEPVLHPVFLCLMLSTDLAASVSTVPKLAIFWCGAGHI				
SASACLAQWFFTHAFOMESTVLLAMAPDYVAICPLRYATILDTIIAHIGVAIV				
PGSLMLPCPLTIGRLNFCQSHVILHTYCEHMAVKLAGDTRPNRVGLTAAALVIG				
VDLFCIGLSYALSAQAVLSSHEARSKALCTCGSHVCVILISYTPALFSEFTFRFGH				
HVPVHIHILLANVILLPPALNPVTVGVTKQIRKRVVRVFGSQGMGKASE"				
Alignment Scores:				
Pred. No.: 1,94e-151				
Length: 1390				
Matches: 1691.00				
Conservative: 1				
Percent Similarity: 99.09%				
Best Local Similarity: 98.78%				
Query Match: 98.43%				
Indels: 0				
Gaps: 0				
DB:				
US-10-081-775-2 (1-329) x AB065534 (1-1390)				
QY 1 MetSerSerThrLeuGlyHisAsnMetGluSerProHisHisThrAspValAspProSer 20				
Db 201 ATGTCCAGCACTTGTGGCACACATGGAATCTCTTAATCACACTGATGTGACCTTCT 260				
QY 21 ValPhePheLeuLeuGlyIleProGlyLeuGluGlnPheHisLeuTrpLeuSerLeuPro 40				
Db 261 GTCTCTTCTCTCTGGGCATCCAGGTCTGGAAACAATTCATTGTGGTCTCTCACTCCCT 320				
QY 41 ValCysGlyLeuGlyThrAlaThrIleValGlyAsnIleThrIleLeuValValAla 60				

JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL

Unpublished
2 (bases 1 to 172027)
Waterston,R.H.
Direct Submission
Submitted (19-MAR-2000) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
On Sep 1, 2000 this sequence version replaced gi:9838305.

COMMENT

----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: <http://genome.wustl.edu/gsc/index.shtml>
----- Project Information -----
Center project name: H.NH0658K18
----- Summary Statistics -----
Sequencing vector: M13; 100%
Sequencing vector: plasmid; 0%
Chemistry: Dye-terminator ET; 100% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 16096 bases at least Q40
Consensus quality: 163968 bases at least Q30
Consensus quality: 165626 bases at least Q20
Insert size: 188000; agarose-fp
Insert size: 171387; sum-of-contigs
Quality coverage: 4.33 in Q20 bases; agarose-fp
Quality coverage: 4.82 in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently
* consists of 19 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 1526: contig of 1526 bp in length
1527 1626: gap of unknown length
1627 2975: contig of 1349 bp in length
2976 3075: gap of unknown length
3076 5074: contig of 1998 bp in length
5074 5173: gap of unknown length
5173 9317: contig of 4144 bp in length
9317 9417: gap of unknown length
9418 13195: contig of 3778 bp in length
13195 13295: gap of unknown length
13295 16482: contig of 3187 bp in length
16482 16582: gap of unknown length
16582 21965: contig of 5383 bp in length
21965 22065: gap of unknown length
22065 33167: contig of 1102 bp in length
33167 33267: gap of unknown length
33267 43162: contig of 9895 bp in length
43162 43262: gap of unknown length
43262 53157: contig of 9895 bp in length
53157 53257: gap of unknown length
53257 53258: contig of 8708 bp in length
53258 61966: gap of unknown length
61966 70847: contig of 8782 bp in length
70847 70948: gap of unknown length
70948 84297: contig of 13350 bp in length
84297 84398: gap of unknown length
84398 100019: contig of 15622 bp in length
100019 100119: gap of unknown length
100119 114785: contig of 14666 bp in length
114785 114885: gap of unknown length
114885 131872: contig of 16987 bp in length
131872 131972: gap of unknown length
131972 147618: contig of 15646 bp in length
147618 171498: gap of unknown length
171498 171598: contig of 23780 bp in length
171598: gap of unknown length

FEATURES
source

* 171599 172027: contig of 429 bp in length.
Location/Qualifiers
1..172027
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="11"
/clone="RP11-658K18"
1..1526
/note="assembly_name:Contig14"
1627..2975
/note="assembly_name:Contig17"
3076..5073
/note="assembly_name:Contig18"
5174..9317
/note="assembly_name:Contig19"
9418..13195
/note="assembly_name:Contig20"
13296..16482
/note="assembly_name:Contig21"
16583..21965
/note="assembly_name:Contig22"
22066..33167
/note="assembly_name:Contig23"
33268..43162
/note="assembly_name:Contig24"
43263..53157
/note="assembly_name:Contig25"
53258..61965
/note="assembly_name:Contig26"
62066..70847
/note="assembly_name:Contig27"
clone_end:T7
vector_side:right
70948..84297
/note="assembly_name:Contig28"
84398..100019
/note="assembly_name:Contig29"
100120..114785
/note="assembly_name:Contig30"
114886..131872
/note="assembly_name:Contig31"
131973..147618
/note="assembly_name:Contig32"
147719..171498
/note="assembly_name:Contig33"
171599..172027
/note="assembly_name:Contig16"

ORIGIN

Alignment Scores:
Pred. No.: 5.29e-149 Length: 172027
Score: 1691.00 Matches: 325
Percent Similarity: 99.09% Conservative: 1
Best Local Similarity: 98.78% Mismatches: 3
Query Match: 98.43% Indels: 0
DB: 2 Gaps: 0

US-10-081-775-2 (1-329) x AC026090 (1-172027)

QY 1 MetSerThrLeuGlyHisAsnMetGluSerProHisHisThrAspValAspProSer 20
89467 ATGTCCAGCACTCTTGCCCAACATGGAATCTCTTAATCACACTGATGTTGACCCCTCT 89408
QY 21 ValPhePheLeuLeuGlyIleProGlyLeuGluGlnPheHisLeuTrpLeuSerLeuPro 40
89407 GTCTTCTCTCTCTGGCCATCCAGGCTCGAACAATTCATTGTGGCTCTCACTCCCT 89348
QY 41 ValCysGlyLeuGlyThrAlaThrIleValGlyAsnIleThrIleLeuValValAla 60
89347 GTGTGGCTTAGGCACAGCCACAAATTTGGGCAATATACTATTCTGTTGTTGTTGCC 89288
QY 61 ThrGluProValLeuHisLysProValTyrLeuPheLeuCysMetLeuSerThrIleAsp 80

Db 89287 ACTGAACAGTCTTGCAACAGCTGTGTACCTTTTCTGTGATGCTCTCAACCATCGAC 89288
Qy 81 LeuAlaAlaSerValSerThrValProLysLeuLeuAlaAlaPheThrCysGlyValaGly 100
Db 89227 TTGGCTGCTCTGCTCCACAGTTCCCAAGCTACTGGCTATCTTCTGGTGGAGCCGGA 89168
Qy 101 HisLeSerAlaSerAlaCysLeuAlaHisMetPhePheLeuHisAlaPheCysMetMet 120
Db 89167 CATATATCTGCTCTGCTGCTGCTGCAAGATGTTCTTCATTCATGCTCTTCGATGATG 89108
Qy 121 GluSerThrValLeuLeuAlaMetAlaPheAspArgTyrValAlaAlaCysHisProLeu 140
Db 89107 GAGTCCACTGTGCTACTGGCAGTCCCTTTGATCGCTAGTGGCCATCTGCCACCATC 89048
Qy 141 ArgTyrAlaThrIleLeuThrAspThrIleAlaHisIleGlyValAlaAlaValVal 160
Db 89047 CGCTATGCCACAATCCTCACTGACACCATATTGCCACATAGGGGTGGCAGCTGTAGTG 88988
Qy 161 ArgGlySerLeuLeuMetLeuProCysProPhePheIleGlyArgLeuAsnPheCysGln 180
Db 88987 CGAGCTCCCTGCTCATGCTCCATGCTCCCTTCTTATTTGGGCGTTTGAACCTTCGCCAA 88928
Qy 181 SerHisValIleLeuHisThrTyrCysGluHisMetAlaValValLysLeuAlaCysGly 200
Db 88927 AGCCATGTGATCTTACACAGCTACTGTGAGCACATGGCTGTGGTGAAGCTGGCTGTGA 88868
Qy 201 AspThrArgProAsnArgValTyrGlyLeuThrAlaAlaLeuLeuValIleGlyValaAsp 220
Db 88867 GACACCGGCTTAACCGTGTATGGGTGACAGCTGACAGCTGCTTGTGTCATTTGGGTTGAC 88808
Qy 221 LeuPheCysIleGlyLeuSerTyrAlaIleLeuAlaGlnAlaValLeuArgLeuSerSer 240
Db 88807 TTGTTTTCATGCTGCTCTCTCATGCCCTTAAGTGACAGCTGCTCTTCTGCTCTCATCC 88748
Qy 241 HisGluAlaArgSerLysAlaLeuGlyThrCysGlySerHisValCysValIleLeuIle 260
Db 88747 CATGAGCTCGTCCAGGCTTACGACCTGTGGTCCCATGCTGTGTGTCATCTCATC 88688
Qy 261 SerTyrThrProAlaLeuPheSerPhePheThrHisArgPheGlyHisHisValProVal 280
Db 88687 TCTTATACACAGCGCTCTCTCTCTTTTACACACCGCTTTGGCCATCAGCTTCAGTFC 88628
Qy 281 HisIleHisIleLeuLeuAlaAsnValTyrLeuLeuLeuProAlaLeuAsnProVal 300
Db 88627 CATATTCATCTCTTTGGCCAAATGTTATCTGCTTTTGGCCACCTGCTTAACTCTGTG 88568
Qy 301 ValTyrGlyValLysThrLysGlnIleArgLysArgValValArgValPheGlnSerGly 320
Db 88567 GTATATGNGTAAACCAACAGATCCGTAAAGAGTTGTGAGGGTGTTCAAAGTGG 88508
Qy 321 GlnGlyMetGlyIleLysAlaSerGlu 329
Db 88507 CAGGGAATGGCATCAAGCATCTCGAG 88481
RESULT 9
AC116156/c
LOCUS Homo sapiens chromosome 11, clone RP11-44D14, complete sequence.
DEFINITION AC116156
ACCESSION AC116156.3 GI:20522208
VERSION
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 176597)
REFERENCE Birren, B., Linton, L., Nusbaum, C. and Lander, E.
TITLE Homo sapiens chromosome 11, clone RP11-44D14
JOURNAL Unpublished
AUTHORS Zainoun, J., Zembek, L., Nusbaum, C., Lander, E., Ali, A., Allen, N.,
Anderson, S., Barna, N., Bastien, V., Bloom, T., Boguslavskiy, L.,
Boukhgalter, B., Brown, A., Camarata, J., Chang, J.,
Cook, A., Cooke, P., DeArelano, K., Dewar, K., Diaz, J.S., Dodge, S.,
Faroo, S., Ferreira, P., FitzHugh, W., Gage, D., Galagan, J., Gardyna, S.,
Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N.,
Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C.,
Kamat, A., Karatas, A., Kells, C., LaRoque, K., Lamazares, R.,
Lander, E., Lehoczy, J., Levine, R., Lindblad-Toh, K., Liu, G.,
MacLean, C., Macdonald, P., Major, J., Marquis, N., Matthews, C.,
McCarthy, M., McEwan, P., McKernan, K., Meldrum, J., Meneus, L.,
Mihova, T., Mienga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R.,
Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D.,
Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V.,
Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P.,
Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schupback, R.,
Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N.,
Strauss, N., Subramanian, A., Talamas, J., Testaye, S., Theodore, J.,
Topham, K., Travers, M., Travis, N., Trigilio, J., Vassiliev, H.,
Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G.,
Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Direct Submission
Submitted (18-APR-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
4 (bases 1 to 176597)
Birren, B., Linton, L., Nusbaum, C., Lander, E., Ali, A., Allen, N.,
Anderson, S., Barna, N., Bastien, V., Bloom, T., Boguslavskiy, L.,
Boukhgalter, B., Brown, A., Camarata, J., Chang, J.,
Cook, A., Cooke, P., DeArelano, K., Dewar, K., Diaz, J.S., Dodge, S.,
Faroo, S., Ferreira, P., FitzHugh, W., Gage, D., Galagan, J., Gardyna, S.,
Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N.,
Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C.,
Kamat, A., Karatas, A., Kells, C., LaRoque, K., Lamazares, R.,
Lander, E., Lehoczy, J., Levine, R., Lindblad-Toh, K., Liu, G.,
MacLean, C., Macdonald, P., Major, J., Marquis, N., Matthews, C.,
McCarthy, M., McEwan, P., McKernan, K., Meldrum, J., Meneus, L.,
Mihova, T., Mienga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R.,
Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D.,
Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V.,
Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P.,
Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schupback, R.,
Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N.,
Strauss, N., Subramanian, A., Talamas, J., Testaye, S., Theodore, J.,
Topham, K., Travers, M., Travis, N., Trigilio, J., Vassiliev, H.,
Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G.,
Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Direct Submission
Submitted (25-MAR-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 176597)
Birren, B., Linton, L., Nusbaum, C., Lander, E., Ali, A., Allen, N.,
Anderson, S., Barna, N., Bastien, V., Bloom, T., Boguslavskiy, L.,
Boukhgalter, B., Brown, A., Camarata, J., Chang, J.,
Cook, A., Cooke, P., DeArelano, K., Dewar, K., Diaz, J.S., Dodge, S.,
Faroo, S., Ferreira, P., FitzHugh, W., Gage, D., Galagan, J., Gardyna, S.,
Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N.,
Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C.,
Kamat, A., Karatas, A., Kells, C., LaRoque, K., Lamazares, R.,
Lander, E., Lehoczy, J., Levine, R., Lindblad-Toh, K., Liu, G.,
MacLean, C., Macdonald, P., Major, J., Marquis, N., Matthews, C.,
McCarthy, M., McEwan, P., McKernan, K., Meldrum, J., Meneus, L.,
Mihova, T., Mienga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R.,
Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D.,
Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V.,
Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P.,
Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schupback, R.,
Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N.,
Strauss, N., Subramanian, A., Talamas, J., Testaye, S., Theodore, J.,
Topham, K., Travers, M., Travis, N., Trigilio, J., Vassiliev, H.,
Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G.,
Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Direct Submission
Submitted (18-APR-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
4 (bases 1 to 176597)
Birren, B., Linton, L., Nusbaum, C., Lander, E., Ali, A., Allen, N.,
Anderson, S., Barna, N., Bastien, V., Bloom, T., Boguslavskiy, L.,
Boukhgalter, B., Brown, A., Camarata, J., Chang, J.,
Cook, A., Cooke, P., DeArelano, K., Dewar, K., Diaz, J.S., Dodge, S.,
Faroo, S., Ferreira, P., FitzHugh, W., Gage, D., Galagan, J., Gardyna, S.,
Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N.,
Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C.,
Kamat, A., Karatas, A., Kells, C., LaRoque, K., Lamazares, R.,
Lander, E., Lehoczy, J., Levine, R., Lindblad-Toh, K., Liu, G.,
MacLean, C., Macdonald, P., Major, J., Marquis, N., Matthews, C.,
McCarthy, M., McEwan, P., McKernan, K., Meldrum, J., Meneus, L.,
Mihova, T., Mienga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R.,
Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D.,
Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V.,
Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P.,
Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schupback, R.,
Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N.,
Strauss, N., Subramanian, A., Talamas, J., Testaye, S., Theodore, J.,
Topham, K., Travers, M., Travis, N., Trigilio, J., Vassiliev, H.,
Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G.,
Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

TITLE

JOURNAL

REFERENCE
AUTHORS

Submitted (10-MAY-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
 5 (bases 1 to 176597)
 Birren, B., Linton, L., Nusbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S., Barna, N., Bastien, V., Bloom, T., Boguslavskiy, L., Boukhgalter, B., Brown, A., Camarata, J., Campopiano, A., Chang, J., Chazaro, B., Choepel, Y., Colangelo, M., Collins, S., Collamore, A., Cook, A., Cooke, P., DeArellano, K., Dewar, K., Diaz, J. S., Dodge, S., Fero, S., Ferreira, P., FitzHugh, W., Gage, D., Galagan, J., Gardyna, S., Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N., Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., Lakoque, K., Lamazares, R., Landers, T., Lehoczy, J., Levine, R., Lindblad-Toh, K., Liu, G., MacLean, C., Macdonald, P., Major, J., Marquis, N., Matthews, C., McCarthy, M., McEwan, P., McKernan, K., Meldrum, J., Meneus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., Norman, C. H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V., Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P., Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schupbach, R., Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N., Strauss, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Travis, N., Trigglio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

TITLE

JOURNAL

COMMENT

Submitted (13-MAY-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
 On May 10, 2002 this sequence version replaced gi:20177753.
 All repeats were identified using RepeatMasker:
 Smit, A.F.A. & Green, P. (1996-1997)
 http://ftp.genome.washington.edu/RM/RepeatMasker.html

----- Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR

Web site: <http://www-seq.wi.mit.edu>

Contact: sequence_submissions@genome.wi.mit.edu

----- Project Information

Center project name: L26176

Center clone name: 44_D_14

FEATURES
source

Location/Qualifiers
 1..176597
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /chromosome="11"
 /map="11"
 /clone="RP11-44D14"
 /clone_libs="RPC1-11 Human Male BAC"

repeat_region

775..802

/rpt_family=" (TTCC)n"

1179..1224

/rpt_family=" (CA)n"

1709..2076

/rpt_family="Charlieb"

2893..2965

/rpt_family="CA-rich"

4667..4718

/rpt_family="GA-rich"

6660..6815

/rpt_family="L2"

complement (6896..7766)

/rpt_family="L1M2"

complement (8463..8683)

/rpt_family="L1M4"

8684..8720

/rpt_family=" (TTTA)n"

complement (8721..9405)

/rpt_family="L1PM4"

complement (9406..9469)

/rpt_family="L1M4"

complement (9486..9561)

repeat_region
 10351..10432
 /rpt_family="FLAM"
 repeat_region
 10833..10896
 /rpt_family="HAL1"
 repeat_region
 complement (12312..12520)
 /rpt_family=" (TA)n"
 repeat_region
 complement (13038..13585)
 /rpt_family="MER77"
 repeat_region
 complement (14764..15303)
 /rpt_family="L1MB5"
 repeat_region
 complement (15304..15618)
 /rpt_family="AluY"
 repeat_region
 complement (15619..16311)
 /rpt_family="L1MB5"
 repeat_region
 complement (16312..16610)
 /rpt_family="AluSc"
 repeat_region
 complement (16611..16666)
 /rpt_family="L1MB5"
 repeat_region
 complement (16677..16816)
 /rpt_family="L1MA6"
 repeat_region
 16815..19052
 /rpt_family="L1PA5"
 repeat_region
 complement (19063..19103)
 /rpt_family="L1MA6"
 repeat_region
 complement (19102..20472)
 /rpt_family="L1PA16"
 repeat_region
 complement (20467..20806)
 /rpt_family="L1MB5"
 repeat_region
 20806..22902
 /rpt_family="L1MB5"
 repeat_region
 complement (22888..22923)
 /rpt_family="L1MA7"
 repeat_region
 complement (22937..23722)
 /rpt_family="L1M4"
 repeat_region
 complement (23871..24202)
 /rpt_family="L1M4"
 repeat_region
 25092..25416

Alignment Scores:

Pred. No.: 5,46e-149 Length: 176597
 Score: 1691.00 Matches: 325
 Percent Similarity: 99.09% Conservative: 1
 Best Local Similarity: 98.78% Mismatches: 3
 Query Match: 98.43% Indels: 0
 DB: 9 Gaps: 0

US-10-081-775-2 (1-329) x AC116156 (1-176597)

Qy 1 MetSerThrLeuGlyHisAsnMetGluSerProHisHisThrAspValAspProSer 20
 Db 159269 ATGTCAGCACTCTTGGCCACACATGGAATCTCTAATACACATGATGTTGACCTTCT 159210
 Qy 21 ValPhePheLeuLeuGlyIleProGlyLeuGluGlnPheHisLeuTrpLeuSerLeuPro 40
 Db 159209 GTCTTCTCTCTCCCTGGGCATCCCAAGGCTCGAACAAATTTCAATTTGGCTCTCACTCCCT 159150
 Qy 41 ValCysGlyLeuGlyThrAlaThrIleValGlyAsnIleThrIleValValValValAla 60
 Db 159149 GTGTGTGGCTTAGGCACAGCCACCAATTTGTGGCAATATACTATTCTGGTGTGTGGCC 159090
 Qy 61 ThrGluProValLeuHisLysProValTrpLeuPheLeuCysMetLeuSerThrIleAsp 80
 Db 159089 ACTGAACCACTGTGCAAGCCTGTGTACCTTTTCTGTGATGCTCTCAACCATCGAC 159030
 Qy 81 LeuAlaAlaSerValSerThrValProLysLeuLeuAlaIlePheTrpCysGlyAlaGly 100
 Db 159029 TTGGCTGCCCTCTGTCTCCACAGTTCCCAAGCTACTTGCTACTTCTCTGGTGGAGCCGA 158970
 Qy 101 HisIleSerAlaSerAlaCysLeuAlaHisMetPhePheIleHisAlaPheCysMetMet 120
 Db 158969 CATATATCGCCTCTGCCTGCTGGCCACAGATGTTCTTCAATTCATGCTTCTTGCATGATG 158910


```
Qy 121 GluSerThrValLeuLeuAlaMetAlaPheAspArgTyrValAlaIleCysHisProLeu 140
Db 158909 GAGTCCACTGTGCTACTGGCCATGCGCTTTGATCGTACGTGGCCATCTGCCACCCACTC 158850

Qy 141 ArgTyrAlaThrIleLeuThrAspThrIleIleAlaHisIleGlyValAlaAlaValVal 160
Db 158849 CGCTATGGCAAACTCTCAGTACACCATCATTTGCCACATAGGGGTGGCAGCTGTAGTG 158790

Qy 161 ArgGlySerLeuLeuMetLeuProCysProPhePheIleGlyArgLeuAsnPheCysGln 180
Db 158789 CGAGGCTCCCTGCTCATGCTCCCATGTCCCTTCTTATTTGGCGCTTTGAACCTTCTGCCAA 158730

Qy 181 SerHisValIleLeuHisThrTyrCysGluHisMetAlaValValLeuLeuAlaCysGly 200
Db 158729 AGCCATGTGATCTTACACACGTACTGTGAGCACATGGCTGTGGTGAAGCTGGCCCTGTGGA 158670

Qy 201 AspThrArgProAsnArgValTyrGlyLeuThrAlaAlaLeuValIleGlyValAsp 220
Db 158669 GACACCGCCCTAACCGTGTGTATGGGTGACAGCTGCACTGTTGTTGCTATTTGGGTTGAC 158610

Qy 221 LeuPheCysIleGlyLeuSerTyrAlaLeuIleAlaGlnAlaValLeuArgLeuSerSer 240
Db 158609 TTTGTTTGCATTGGTCTCTCTATGCCCTAAGTGCACAAAGCTGCTCTCGCCTCTCATCC 158550

Qy 241 HisGluAlaArgSerIleAlaLeuGlyThrCysGlySerHisValCysValIleLeuIle 260
Db 158549 CATGAAGCTCGTCCAAAGCCCTTAGGACCTGTGGTTCCTCATCTGTGTATCTCATC 158490

Qy 261 SerTyrThrProAlaLeuPheSerPhePheThrHisArgPheGlyHisValProVal 280
Db 158489 TCTTATACACAGCCCTCTCTCTCTTTTACACACCGCTTTGGCCATCAGCTTCCAGTC 158430

Qy 281 HisIleHisIleLeuLeuAlaAsnValTyrLeuLeuLeuProAlaLeuAsnProVal 300
Db 158429 CATATTACATTCTTTGGCCAAATGTTATCTGCTTTTGGCCACCTGCTCTTAATCTGTG 158370

Qy 301 ValTyrGlyValIleThrIleArgIleArgValValArgValPheGlnSerGly 320
Db 158369 GTATATGAGATTAAACCAAAAGATCCGTAAGAGATTGTTCAGGCTGTTTCAAAGTGGG 158310

Qy 321 GlnGlyMetGlyIleLeuAlaSerGlu 329
Db 158309 CAGGGAATGGGCATCAAGGCATCTGAG 158283

RESULT 10
AX241675 963 bp DNA linear PAT 26-SEP-2001
LOCUS Sequence 423 from Patent WO0127158.
DEFINITION AX241675
ACCESSION AX241675
VERSION AX241675.1 GI:15798550
KEYWORDS synthetic construct
SOURCE synthetic construct
ORGANISM artificial sequences.
REFERENCE 1
AUTHORS Bellenson,J., Smith,D., Lancet,D., Glusman,G., Fuchs,T. and Yanai,I.
TITLE Olfactory receptor sequences
JOURNAL Patent: WO 0127158-A 423 19-APR-2001;
Digiscents (US) ; YEDA RESEARCH AND DEVELOPMENT COMPANY, LTD. (IL)
FEATURES
source
1..963
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note=" (H389272 nucleotide) "
```

Alignment Scores: 1.26e-147 Length: 963
Pred. No.: 317
Score: 1649.00
Percent Similarity: 99.07%
Best Local Similarity: 98.75%
Conservative: 1
Mismatch: 3

```
Query Match: 95.98% Indels: 0
DB: 6 Gaps: 0
US-10-081-775-2 (1-329) x AX241675 (1-963)

Qy 9 MetGluSerProHisHisThrAspValAspProSerValPhePheLeuLeuGlyIlePro 28
Db 1 ATGGAAATCTCTAATCACACTGATGTTGACCTTCTGTCTTCTTCTCTCTCTCTCTCTCT 60

Qy 29 GlyLeuGlnPheHisLeuThrLeuSerLeuProValCysGlyLeuGlyThrAlaThr 48
Db 61 GGTCTGGAAACAATTTCAATTTGTGCTCTCACTCCCTGTGTGTGGCTTAGGCACAGCCACA 120

Qy 49 IleValGlyAenIleThrIleLeuValValAlaThrGluProValLeuHisIlePro 68
Db 121 ATTTGGGCAATAATAACTATTCTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 180

Qy 69 ValTyrLeuPheLeuCysMetLeuSerThrIleAspLeuAlaAlaSerValSerThrVal 88
Db 181 GTGTACCTTTTCTGTGTCATGCTCTCAACCATCGACTTGGCTTGGCTTGTCTTCTCCACAGT 240

Qy 89 ProLysLeuLeuAlaIlePheThrCysGlyAlaGlyHisIleSerAlaSerAlaCysLeu 108
Db 241 CCCAAGCTTACTGGCTATCTTCTGTGTGGAGCCGACATATATCTGCTTGCCTTGCCTTGCCTG 300

Qy 109 AlaHisMetPhePheIleHisAlaPheCysMetMetGluSerThrValLeuLeuAlaMet 128
Db 301 GCACAGATGTTCTTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCAT 360

Qy 129 AlaPheAspArgTyrValAlaIleCysHisProLeuArgTyrAlaThrIleLeuThrAsp 148
Db 361 GCCTTTGATCGTACGTGGCCATCTGCCACCACTCCGCTATGCCACCAATCTCTCACTGAC 420

Qy 149 ThrIleIleAlaHisIleGlyValAlaValAlaValArgGlySerLeuLeuMetLeuPro 168
Db 421 ACCATCATTTGGCCACATAGGGGTGGCAGCTGTAGTGGAGGCTCCCTGCTCATGCTCCCA 480

Qy 169 CysProPhePheIleGlyArgLeuAsnPheCysGlnSerHisValIleLeuHisThrTyr 188
Db 481 TGTCTCTCTCTTATTTGGCGTTTGAACCTTCTGCCAAAGCCATGTGATCTCTACACAGTAC 540

Qy 189 CysGluHisMetAlaValValLysLeuAlaCysGlyAspThrArgProAsnArgValTyr 208
Db 541 TGTGAGCACATGGCTGTGTGAAGCTGGCTGTGGAGACACCCAGGCTTAAACCGTGTGTAT 600

Qy 209 GlyLeuThrAlaAlaLeuValIleGlyValAlaAspLeuPheCysIleGlyLeuSerTyr 228
Db 601 GGGCTGACAGCTGCACCTGTGTGTCATTTGGGGTTGACTTGTTTTGTGATTTGTTCTTCTAT 560

Qy 229 AlaLeuIleAlaGlnAlaValLeuArgLeuSerSerHisGluAlaArgSerLysAlaLeu 248
Db 661 GCCCTAAGTGCAAGCTGCTCTTCCCTCTCATCCATGAAGCTCGGTCCAAGGCCCTA 720

Qy 249 GlyThrCysGlySerHisValCysValIleLeuIleSerTyrThrProAlaLeuPheSer 268
Db 721 GGGACCTGTGGTTCCCATGCTGTGTGTCATCTCTTATACACACCCCTCTTCTCTCC 780

Qy 269 PhePheThrHisArgPheGlyHisValProValHisIleHisIleLeuLeuAlaAsn 288
Db 781 TTTTATTACACACCGCTTTGGCCATCAGTTCCAGTCCATATTACATTTCTTTTGGCCAT 840

Qy 289 ValTyrLeuLeuLeuProProAlaLeuAsnProValValTyrGlyValLysThrLysGln 308
Db 841 GTTATCTGCTTTTGGCCACTGCTCTTAATCTGTGTATATGGAGTTAAGACCAACAG 900

Qy 309 IleArgLysArgValValArgValPheGlnSerGlyGlnGlyMetGlyIleLysAlaSer 328
Db 901 ATCCGTAAGAGATTGTGAGGGTGTTCAGAGTGGGAGGAGTGGGATCGGATCTCT 960

Qy 329 Glu 329
Db 961 GAG 963
```


US-10-081-775-2 (1-329) x AX448757 (1-966)

```
Qy 9 MetGluSerProHisHisThrAspValAspProSerValPhePheLeuLeuGlyIlePro 28
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1 ATGGAATCTCTAATACACTGATGTTGACCGCTTCTGTCTTCTCTCTCTGGCATCCCA 60

Qy 29 GlyLeuGluGlnPheHisLeuTrpLeuSerLeuProValCysGlyLeuGlyThrAlaThr 48
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
61 GGCTCGGAACAATTTCAITTTGGCTCTCACTCCCTGTGTGGCTTAGGCACAGCCACA 120

Qy 49 IleValGlyAenIleThrIleLeuValValAlaThrGluProValLeuHisIysPro 68
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
121 ATTGTGGGCAATAAATACTATTCTGTGTGTGTGCACTGAACAGCTTGTGCACAAGCCT 180

Qy 69 ValTyrLeuPheLeuCysMetLeuSerThrIleAspLeuAlaAlaSerValSerThrVal 88
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
181 GTGTACCTTTTCTGTGTGATGCTCTCAACCATCGACTTGGCTGTCTGTCTCCACAGTT 240

Qy 89 ProLysLeuLeuAlaIlePheTrpCysGlyAlaGlyHisIleSerAlaSerAlaCysLeu 108
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
241 CCCAAGCTACTGGCTATCTTCTGTGTGTGGAGCCGACATATATCTGCTCTGCTGCTG 300

Qy 109 AlaHisMetPhePheIleHisAlaPheCysMetMetGluSerThrValLeuLeuAlaMet 128
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
301 GCACAGATGTTCTTCTCATCTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 360

Qy 129 AlaPheAspArgTyrValAlaIleCysHisProLeuArgTyrAlaThrIleLeuThrAsp 148
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
361 GCCTTGTATCGCTAGCTGCGCATCTGCCACCATCTGCCCATCTGCCCATCTGCCCAT 420

Qy 149 ThrIleIleAlaHisIleGlyValAlaValAlaValArgGlySerLeuLeuMetLeuPro 168
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
421 ACCATCATGCGCCACATAGGCGTGCAGCTGTAGTGCAGGCTCCTGCTCATGCTCCCA 480

Qy 169 CysProPhePheIleGlyArgLeuAsnPheCysGlnSerHisValIleLeuHisThrTyr 188
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
481 TGTCCCTCTTATTGGGCGTTTGAATCTTGCCAAAGCCATGTGATCCTACACAGTAC 540

Qy 189 CysGluHisMetAlaValValIlysLeuAlaCysGlyAspThrArgProAsnArgValTyr 208
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
541 TGTGAGCACATGGCTGTGTGTAAGCTGGCTGTGTGAGACACACAGCCCTTAACCGTGT 600

Qy 209 GlyLeuThrAlaAlaLeuValIleGlyValAspLeuPheCysIleGlyLeuSerTyr 228
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
601 GGGCTGACAGCTGCACTGTGTGCTCATTTGGGTTGACTGTTTGGATTTGGTCTCTCTAT 660

Qy 229 AlaLeuIleAlaGlnAlaValLeuArgLeuSerSerHisGluAlaArgSerIlysAlaLeu 248
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
661 GCCCTAAGTGACAAAGCTGCTTCTGCTCTCATCCCATGAAGCTCGGTCCAAGGCCCTA 720

Qy 249 GlyThrCysGlySerHisValCysValIleLeuIleSerTyrThrProAlaLeuPheSer 268
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
721 GGGACCTGTGGTTCCTCATGTGTGTATCCCTCATCTCTTATACACAGCCCTCTTCTCC 780

Qy 269 PhePheThrHisArgPheGlyHisHisValProValHisIleHisIleLeuLeuAlaAsn 288
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
781 TTTTITACACACCGCTTTGGCCATCAGCTTCAGTCCATATTCACATCTTTTGGCCAT 840

Qy 289 ValTyrLeuLeuLeuProProAlaLeuAsnProValValTyrGlyValIlysThrIysGln 308
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
841 GTTTATCTGCTTTTGGCCACCTCTCTTATCTGTGTATATGAGTGTAAAGACCAACAG 900

Qy 309 IleArgLysArgValValArgValPheGlnSerGlyGlnGlyMetGlyIleIlysAlaSer 328
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
901 ATCCGTAAAGAGTTGTGAGGTGTGTTTCAAGTGGCGAGGGAATGGGCATCAAGGCACT 960

Qy 329 Glu 329
Db |||||
961 GAG 963
```

RESULT 13

AC109341/c

LOCUS

202761 bp

DNA

linear

PRI 14-MAY-2002

DEFINITION
AC109341
VERSION
KEYWORDS
SOURCE
ORGANISM

Homo sapiens chromosome 11, clone CTD-2504M7, complete sequence.
AC109341.7 GI:20564454
HTG.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 202761)
Birren, B., Linton, L., Nusbaum, C. and Lander, E.
Homo sapiens chromosome 11, clone CTD-2504M7
Unpublished
2 (bases 1 to 202761)

REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS

Birren, B., Linton, L., Nusbaum, C., Lander, E., Ali, A., Allen, N.,
Anderson, S., Barna, N., Bastien, V., Boguslavskiy, L., Boukhgalter, B.,
Brown, A., Camarata, J., Campopiano, A., Chang, J., Chazaro, B.,
Choepel, Y., Colangelo, M., Collins, S., Collymore, A., Cook, A.,
Cooke, P., DeArelano, K., Dewar, K., Diaz, J.S., Dodge, S., Faro, S.,
Ferreira, P., FitzHugh, W., Gage, D., Galagan, J., Gardyna, S.,
Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N.,
Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C.,
Kamat, A., Karats, A., Kells, C., LaRocque, K., Lamazares, R.,
Landers, T., Lehoczy, J., Levine, R., Liu, G., MacLean, C.,
Macdonald, P., Major, J., Marquis, N., Matthews, C., McCarthy, M.,
McEwan, P., McKernan, K., Meldrim, J., Meneus, L., Mihova, T.,
Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C.,
Norman, C.H., O'Connor, T., O'Donnell, P., O'Neil, D., Oliver, J.,
Peterson, K., Phunkhang, P., Pierre, N., Pollara, V., Raymond, C.,
Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P., Roman, J.,
Rosetti, M., Roy, A., Santos, R., Schauer, S., Schupback, R., Seaman, S.,
Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N.,
Strauss, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J.,
Topham, K., Travers, M., Travis, N., Trigilio, J., Vassiliev, H.,
Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G.,
Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Direct Submission

Submitted (03-FEB-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA

3 (bases 1 to 202761)

TITLE
JOURNAL
REFERENCE
AUTHORS

Birren, B., Linton, L., Nusbaum, C., Lander, E., Ali, A., Allen, N.,
Anderson, S., Barna, N., Bastien, V., Bloom, T., Boguslavskiy, L.,
Boukhgalter, B., Brown, A., Camarata, J., Campopiano, A., Chang, J.,
Chazaro, B., Choepel, Y., Colangelo, M., Collins, S., Collymore, A.,
Cooke, A., Cooke, P., DeArelano, K., Dewar, K., Diaz, J.S., Dodge, S.,
Faro, S., Ferreira, P., FitzHugh, W., Gage, D., Galagan, J., Gardyna, S.,
Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N.,
Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C.,
Kamat, A., Karats, A., Kells, C., LaRocque, K., Lamazares, R.,
Landers, T., Lehoczy, J., Levine, R., Lindblad-Toh, K., Liu, G.,
MacLean, C., Macdonald, P., Major, J., Marquis, N., Matthews, C.,
McCarthy, M., McEwan, P., McKernan, K., Meldrim, J., Meneus, L.,
Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R.,
Norman, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neil, D.,
Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V.,
Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P.,
Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schupback, R.,
Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N.,
Strauss, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J.,
Topham, K., Travers, M., Travis, N., Trigilio, J., Vassiliev, H.,
Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G.,
Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Direct Submission

Submitted (03-MAY-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA

4 (bases 1 to 202761)

TITLE
JOURNAL
REFERENCE
AUTHORS

Birren, B., Linton, L., Nusbaum, C., Lander, E., Ali, A., Allen, N.,
Anderson, S., Barna, N., Bastien, V., Bloom, T., Boguslavskiy, L.,
Boukhgalter, B., Brown, A., Camarata, J., Campopiano, A., Chang, J.,
Chazaro, B., Choepel, Y., Colangelo, M., Collins, S., Collymore, A.,
Cooke, A., Cooke, P., DeArelano, K., Dewar, K., Diaz, J.S., Dodge, S.,
Faro, S., Ferreira, P., FitzHugh, W., Gage, D., Galagan, J., Gardyna, S.,
Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N.,
Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C.,


```
Db 69511 TGTGTGGCTTGGACAGCCACAAATGTGGCAATATACTATTCGGTGTGTGGC 69452
Qy 60 aThrGluProValLeuHisLysProValTyrLeuPheLeuCysMetLeuSerThrIleAs 80
Db 69451 CACTGAACCACTCTTGACAGACCTGTGTACTCTTTCTTGTCATCTCTCAACCATCGA 69392
Qy 80 pLeuAlaSerValSerThrValProLysLeuLeuAlaIlePheTyrCysGlyAlaG1 100
Db 69391 CTTGGCTGCTCTGTCTCCACAGTTCCTCCAGCTACTGGCTATCTCTGTGTGGAGCCGG 69332
Qy 100 yHisIleSerAlaSerAlaCysLeuAlaHisMetPhePheIleHisAlaPheCysMetMe 120
Db 69331 ACATATATCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 69272
Qy 120 tGluSerThrValLeuLeuAlaMetAlaPheAspArgTyrValAlaIleCysHisProLe 140
Db 69271 GGAGTCCACTGTGCTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 69212
Qy 140 uArgTyrAlaThrIleLeuThrAspThrIleIleAlaHisIleGlyValAlaAlaValVa 160
Db 69211 CCGCTATGCCACAAATCCTCACTGACACCATCATTCGCCACATAGGGGTGGCAGCTGTAGT 69152
Qy 160 lArgGlySerLeuLeuMetLeuProCysProPhePheIleGlyArgLeuAsnPheCysG1 180
Db 69151 GCGAGGCTCCCTGCTCAIGCTCCCAATGCCCTCTTTATTTGGGCGTTTGAACCTCTGCCA 69092
Qy 180 nSerHisValIleLeuHisThrTyrCysGluHisMetAlaValValLysLeuAlaCysG1 200
Db 69091 AAGCCATGTGATCCTACACACATCTGTAC----ATGGCTGTGTGTGAAGCTGGCCTGTGG 69036
Qy 200 yAspThrArgProAsnArgValTyrGlyLeuThrAlaAlaLeuValIleGlyValAs 220
Db 69035 AGACACCAAGGCTAAACCGTGTGTATGGGCTGACACCTGCTGCTGCTGCTGCTGCTGCTG 68976
Qy 220 pLeuPheCysIleGlyLeuSerTyrAlaLeuIleAlaGlnAlaValLeuArgLeuSerSe 240
Db 68975 CTTGTTTTCATGTGTCTCTCTATGCTTATGCTTATGCTTATGCTTATGCTTATGCTTATG 68916
Qy 240 rHisGluAlaArgSerLysAlaLeuGlyThrCysGlySerHisValCysValIleLeu11 260
Db 68915 CCATGAAGCTGGTCCAGGCTTCCAGGCTTCCAGGCTTCCAGGCTTCCAGGCTTCCAGGCT 68856
Qy 260 eSerTyrThrProAlaLeuPheSerPhePheThrHisArgPheGlyHisHisValProVa 280
Db 68855 CTCCTTATACACAGCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 68796
Qy 280 HisIleHisIleLeuLeuAlaAsnValTyrLeuLeuLeuProProAlaLeuAsnProVa 300
Db 68795 CCATATTCATCTCTTTGGCCAAATGTTTATCTGCTTTTGGCCAACTGCTCTTAATCTGT 68736
Qy 300 lValTyrGlyValLysThrLysGlnIleArgLysArgValValValArgValPheGlnSerG1 320
Db 68735 GGTATATGAGTAAAGCAACACAGATCCGTAAAGAGTTGTCAAGGTGTTTCAAGTGG 68676
Qy 320 yGlnGlyMetGlyIleLysAlaSerGlu 329
Db 68675 GCAGAAATGGGCATCAAGGCATCTGAG 68648

RESULT 14
AC147586
LOCUS AC147586 181386 bp DNA linear HTG 18-DEC-2003
DEFINITION Otalemur garnettii clone CH256-553F22, WORKING DRAFT SEQUENCE.
ACCESSION AC147586
VERSION AC147586.1 GI:40018680
KEYWORDS HTG; HTGS PHASE2; HTGS DRAFT.
SOURCE Otalemur garnettii (small-eared galago)
ORGANISM Otalemur garnettii
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Screechhorns; Galgonidae; Otalemur.
REFERENCE 1 (bases 1 to 181386)
AUTHORS Antonellis.A., Ayele.K., Benjamin.B., Blakesley.R.W.,
Bouffard.G.G., Brinkley.C., Brooks.S., Chu.G., Coleman.B.,
```

TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

Coleman,H., Daki,N., Engle,J., Granite,S., Guan,X., Gupta,J.,
Haghighi,P., Han,J., Hansen,N., Ho,S.-L., Hu,P., Hurle,B.,
Idol,W.R., Jones,C., Karlins,E., Kim,H., Kwong,P., Laric,P.,
Larson,S., Lee-Jin,S.-Q., Legaspi,R., Maduro,Q.L., Maduro,V.B.,
Margulies,E.H., Masiello,C., Maskeri,B., McDowell,J.,
Mullikin,J.C., Paguirigan,C., Portnoy,M.E., Prasad,A., Puri,O.,
Reddi-Dugue,N., Schandler,K., Schueler,M.G., Shah,K., Sison,C.,
Stantipop,S., Thomas,J.W., Thomas,P.J., Tsipouri,V., Vogt,J.L.,
Wecherby,K.D., Young,A. and Green,E.D.
NISC Comparative Sequencing Initiative
Unpublished
2 (bases 1 to 181386)
Green,E.D.
Direct Submission
Submitted (18-DEC-2003) NIH Intramural Sequencing Center, 8717
Gromont Circle, Gaithersburg, MD 20877, USA
----- Genome Center
Center: NIH Intramural Sequencing Center
Web site: http://www.nisc.nih.gov
Contact: nisc.zoo@hri.nih.gov
----- Project Information
Center project name: f11
Center clone name: 553F22
----- Summary Statistics
Sequencing vector: plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 180937 bases at least Q40
Consensus quality: 181081 bases at least Q30
Consensus quality: 181163 bases at least Q20
Insert size: 177000; agarose-fp
Insert size: 181386; sum-of-ctgigs
Quality coverage: 13.31x in Q20 bases; agarose-fp
Quality coverage: 12.99x in Q20 bases; sum-of-ctgigs

* NOTE: This is a 'working draft' sequence. It currently
* consists of 1 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submitter.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
* 1 181386: contig of 181386 bp in length.

FEATURES
source

1..181386
/organism="Otalemur garnettii"
/mol_type="genomic DNA"
/db_xref="taxon:30611"
/clone="CH256-553F22"
/note="BAC resource: http://bacpac.chori.org/"
misc_feature
1..181386
/note="assembly_fragment
clone_end:SP6
vector_side:right"

ORIGIN

Alignment Scores:

Pred. No.: 2,6e-136 Length: 181386
Score: 1558.00 Matches: 297
Percent Similarity: 94.22% Conservative: 13
Best Local Similarity: 90.27% Mismatches: 19
Query Match: 90.69% Indels: 0
DB: Gaps: 0

US-10-081-775-2 (1-329) x AC147586 (1-181386)

Qy 1 MetSerSerThrLeuGlyHisAsnMetGluSerProHisHisThrAspValAspProSer 20

Db 3117 ATATCCGGTACTCTTGGCCAGATCATGAAATCTCTTAATCACACACATGGACCTTCT 3176

```

QY 21 ValPhePheLeuLeuGlyLeuProGlyLeuGluGlnPheHisLeuTrpLeuSerLeuPro 40
Db ::::|||||
3177 ATTTTCTCTCTTGGGTATTCAGGTCTAGAAAAATTTCACTGTGGCTCTCACTTCT 3236

QY 41 ValCysGlyLeuGlyThrAlaThrIleValGlyAsnIleThrIleLeuValValAla 60
Db :|||:|||||
3237 GTGTGCTGCTGGGCACAGCCAGTTGTGGCAACATAACCATCTCTGTTGTGTGGCC 3296

QY 61 ThrGluProValLeuHisLysProValTyrLeuPheLeuCysMetLeuSerThrIleAsp 80
Db :|||||
3297 ACTGAGCCAGCCCTGCACAGGCTGTGTACTTTTCTGTGCAATGTTGTCAACCATCGAC 3356

QY 81 LeuAlaAlaSerValSerThrValProLysLeuLeuAlaIlePheThrCysGlyAlaGly 100
Db :|||||
3357 TTGGCTGCTCTGTCTCCACAGTTCCCAAGCTACTGGGCATTTCTCTGGTGTGGAGCTGA 3416

QY 101 HisIleSerAlaSerAlaCysLeuAlaHisMetPheIleHisAlaPheCysMetMet 120
Db :|||||
3417 CATATATCTGCTCTGCTTGCCTAGCACAGATGTTCTTCATTTCATGCTTCTGCATGATG 3476

QY 121 GluSerThrValLeuLeuAlaMetAlaPheAspArgTyrValAlaIleCysHisProLeu 140
Db :|||||
3477 GAGTCCATCTGCTGCTGCCATGCGCTTGTATCGCTATGTGCCATCTGCCATCCACTC 3536

QY 141 ArgTyrAlaThrIleLeuThrAspThrIleAlaHisIleGlyValAlaValVal 160
Db :|||||
3537 CGCTATGCTACTATCTCTCAGTACCATATTATGCTCGATGGAGTGTGATGTTGGT 3596

QY 161 ArgGlySerLeuLeuMetLeuProCysProPheIleGlyArgLeuAsnPheCysGln 180
Db :|||||
3597 CGAGGCTCCCTGCTATGCTCCCATGCTCCCTTCTTATTTGGCGTTTGAGCTTCTGCCAG 3656

QY 181 SerHisValIleLeuHisThrTyrCysGluHisMetAlaValValLysLeuAlaCysGly 200
Db :|||||
3657 AGCCACGTGATCTGCACACATACGTGAGCAGATGCGCGTTGTGAAGCTGGCTGTGGA 3716

QY 201 AspThrArgProAsnArgValTyrGlyLeuThrAlaAlaLeuLeuValIleGlyValAsp 220
Db :|||||
3717 GACACACAGCCCTAATCGTGTGTCAGGGCTCACAGCTGCACATGTTGGTTCATCGGGTGGAT 3776

QY 221 LeuPheCysIleGlyLeuSerTyrAlaLeuIleAlaGlnAlaValLeuArgLeuSerSer 240
Db :|||||
3777 TTGTTCTGCAATAGTCTCTCTATGCTCTCATGTGTACAAAGCTGCTTTCGCTCTCATCC 3836

QY 241 HisGluAlaArgSerLysAlaLeuGlyThrCysGlySerHisValCysValIleLeuIle 260
Db :|||||
3837 CATGAAGCTGTTCCAGGCTCAGGACCTAGGACCTGTGTTCCCATGCTGTGTATCTCATC 3896

QY 261 SerTyrThrProAlaLeuPheSerPhePheThrHisArgPheGlyHisValProVal 280
Db :|||||
3897 TCTTACACACAGCTCTCTTTCTTTTACATGCTTGGCCACCATGTTCCACTC 3956

QY 281 HisIleHisIleLeuLeuAlaAsnValTyrLeuLeuLeuProProAlaLeuAsnProVal 300
Db :|||||
3957 CATATTCACATTTTGGCCAAATGTTTATCTGCTCTTCCACCTGCTCTTAAACCTGTA 4016

QY 301 ValTyrGlyValLysThrLysGlnIleArgLysArgValValArgValPheGlnSerGly 320
Db :|||||
4017 GTATATGGGTTAAGACCAAGAGATTCGTGAAGGGTTGTGACAGTGTTCGAAGAGGA 4076

QY 321 GlnGlyMetGlyIleLysAlaSerGlu 329
Db :|||||
4077 CAGGGAATGGAATCAAGGCATCTGAG 4103

RESULT 15
AC146635/c
LOCUS
DEFINITION AC146635 239335 bp DNA linear HTG 17-OCT-2003
Otolemur garnettii clone CH256-105A13, WORKING DRAFT SEQUENCE, 4
ordered pieces.
ACCESSION AC146635
VERSION AC146635.2 GI:37700288
KEYWORDS HTG; HTGS_PHASE2; HTGS_DRAFT.

```

SOURCE
ORGANISMREFERENCE
AUTHORS

Otolemur garnettii (small-eared galago)
Otolemur garnettii
Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Chordata; Strepsirhini; Galagonidae; Otolemur.
1 (bases 1 to 239335)
Antonellis, A., Ayele, K., Benjamin, B., Blakesley, R. W.,
Bouffard, G. G., Brinkley, C., Brooks, S., Chu, G., Coleman, B.,
Coleman, H., Engle, J., Granite, S., Guan, X., Gupta, J., Haghighi, P.,
Han, J., Hansen, N., Ho, S.-L., Hu, P., Hurle, B., Idol, J. R.,
Karls, E., Kwong, P., Laric, P., Larson, S., Lee-Lin, S.-O.,
Legaspi, R., Maduro, Q. L., Maduro, V. B., Margulies, E. H., Masiello, C.,
Maskeri, B., McDowell, J., Mullikin, J. C., Paguirigan, K.,
Portnoy, M. E., Prasad, A., Reddix-Dugue, N., Schandler, K.,
Schueler, M. G., Shah, K., Sison, C., Stantrop, S., Thomas, J. W.,
Thomas, P. J., Tsipouri, V., Vogt, J. L., Wetherby, K. D., Young, A. and
Green, E. D.

NISC Comparative Sequencing Initiative

Unpublished
2 (bases 1 to 239335)

Green, E. D.

Direct Submission

Submitted (11-SEP-2003) NIH Intramural Sequencing Center, 8717

Grovermont Circle, Gaithersburg, MD 20877, USA

3 (bases 1 to 239335)

Green, E. D.

Direct Submission

Submitted (17-OCT-2003) NIH Intramural Sequencing Center, 8717

Grovermont Circle, Gaithersburg, MD 20877, USA

On Oct 17, 2003 this sequence version replaced gi:34576313.

----- Genome Center

Center: NIH Intramural Sequencing Center

Center code: NISC

Web site: <http://www.nisc.nih.gov>

Contact: nisc.zoo@hghri.nih.gov

----- Project Information

Center project name: fgs

Center clone name: 105A13

The sequence data in this record represents an 'enhanced' version of a Phase 2 submission. Specifically, the indicated order and orientation of each sequence contig has been established using one or more of the following: read-pair data from individual subclones, overlaps with neighboring clones, alignment with available reference sequence (e.g., human), and/or confirmation by PCR testing. In addition, the sequence assembly is based on at least 8X average coverage in Q20 bases and has been reviewed to rule out gross misassemblies, the low-quality ends of sequence contigs have been trimmed away, and each base is associated with a Phrap-derived quality score.

----- Summary Statistics

Sequencing vector: plasmid; n/a; 100% of reads

Chemistry: Dye-terminator Big Dye; 100% of reads

Assembly program: Phrap; version 0.990319

Consensus quality: 238379 bases at least Q40

Consensus quality: 238694 bases at least Q30

Consensus quality: 238929 bases at least Q20

Insert size: 239000; agarose-fp

Insert size: 239035; sum-of-contigs

Quality coverage: 10.78x in Q20 bases; agarose-fp

Quality coverage: 10.78x in Q20 bases; sum-of-contigs

- * NOTE: This is a 'working draft' sequence. It currently consists of 4 contigs. Gaps between the contigs
- * are represented as runs of N. The order of the pieces
- * is believed to be correct as given, however the sizes
- * of the gaps between them are based on estimates that have
- * provided by the submitter.
- * This sequence will be replaced
- * by the finished sequence as soon as it is available and
- * the accession number will be preserved.
- * 1 4357: contig of 4357 bp in length
- * 4358 4457: gap of unknown length

*	4458	66663: contig of 62206 bp in length
*	66664	66763: gap of unknown length
*	66764	232439: contig of 165676 bp in length
*	232440	232539: gap of unknown length
*	232540	239335: contig of 6796 bp in length.

FEATURES

```

source      1. .239335
            /organism="Otoleumur garnettii"
            /mol_type="genomic DNA"
            /db_xref="taxon:30611"
            /clone="CH256-105A13"
            /clone_lib="CH256"
            1. .4357
            /notes="assembly_fragment
            clone end:T7
            vector_side:left"
            4558. -66663
            /notes="assembly_fragment"
            66764. -232439
            /notes="assembly_fragment"
            232540. -239335
            /notes="assembly_fragment
            clone end:SP6
            vector_side:right"
            misc_feature
            misc_feature
            misc_feature
            misc_feature

```

ORIGIN

Alignment Scores:		
Pred. No.:	3 59e-136	Length: 239335
Score:	158.00	Matches: 297
Percent Similarity:	94.22%	Conservative: 13
Best Local Similarity:	90.21%	Mismatches: 19
Query Match:	90.69%	Indels: 0
DB:	2	Gaps: 0

US-10-081-775-2 (1-329) X AC146635 (1-239335)

Qy	1	MetSerSerThrLeuGlyHisAsnMetGluSerProHisHisThrAspValAspProSer	20
Db	9997	ATATCCGGTACTCTTGGCCAGATCATGAATCTCTTAATACACACACATGGGACCCCTCT	9938
Qy	21	ValPhePheLeuGlyIleProGlyLeuGluGlnPheHisLeuTrpLeuSerLeuPro	40
Db	9937	ATTTTCTTCCTCTGGGTATTCAGGCTAGAAAATTTCACTCTGGCTCTCACTTCT	9878
Qy	41	ValCysGlyLeuGlyThrAlaThrIleValGlyAsnIleThrIleLeuValValAla	60
Db	9877	GTGTGCTGCCTGGGCACAGCCACAGTGTGGGCAACATAACCATCTCTGGTTGTGGCC	9818
Qy	61	ThrGluProValLeuHisLysProValTyrLeuPheLeuCysMetLeuSerThrIleAsp	80
Db	9817	ACTGAGCCAGCCCTGCACAAAGCCTGTGTACTTTCTGTGCATGTGTCAACCAATCGAC	9758
Qy	81	LeuAlaAlaSerValSerThrValProLysLeuLeuAlaIlePheTrpCysGlyAlaGly	100
Db	9757	TTGGCTGCCTCTGTCTCCACAGTTCCCAAGCTACTGGGCATTCTCTGTGTGGAGCTGGA	9698
Qy	101	HisIleSerAlaSerAlaCysLeuAlaHisMetPhePheIleHisAlaPheCysMetMet	120
Db	9697	CATATATCTGCCTCTGCTTGCCTAGCACAGATGTTCTTCAATTATGCCTTCTGCATGATG	9638
Qy	121	GluSerThrValLeuLeuAlaMetAlaPheAspArgTyrValAlaIleCysHisProLeu	140
Db	9637	GAGTCCACTGTGCTGCTGGCCATGGCCTTTGATCGCTATGTGGCCATCTGCCATCCACTC	9578
Qy	141	ArgTyrAlaThrIleLeuThrAspThrIleIleAlaHisIleGlyValAlaAlaValVal	160
Db	9577	CGCTATGTACTATCTCTCACTGACACACATTATTGCTCGCATTTGGAGTGGTAGCTATGGTG	9518
Qy	161	ArgGlySerLeuLeuMetLeuProCysProPhePheIleGlyArgLeuAsnPheCysGln	180
Db	9517	CGAGGCTCCCTGCTATGCTCCCCAGTCCCTTCTTATTTGGCGTTTAGAGCTTCTGCCAG	9458
Qy	181	SerHisValIleLeuHisHisThrTyrCysGluHisMetAlaValValLysLeuAlaCysGly	200

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: August 27, 2004, 18:39:29 ; Search time 548 Seconds

(without alignments)

2954.652 Million cell updates/sec

Title: US-10-081-775-2

Perfect score: 1718

Sequence:

1 MSSTLGHNMESPHHTDVPD.....RKRVRVVFQSGQGMKASE 329

Scoring table:

BLOSUM62

Xgapop 10.0 , Xgapext 0.5

Ygapop 10.0 , Ygapext 0.5

Fgapop 6.0 , Fgapext 7.0

Delop 6.0 , Delext 7.0

Searched: 3237270 seqs, 2460713050 residues

Total number of hits satisfying chosen parameters: 6474540

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlh

-DB=PublishedApplications NA -Q=FASTA -SUFFIX=rnpb -MINMATCH=0.1

-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blom62

-TRANS=human40.cdi -LIST=45 -DOCLIGN=200 -THR SCORE=pct -THR MAX=100

-MAXLEN=2000000000 -USER=US10081775 @CNG1.1.480 @runat 20082004 170214 11184

-NCPU=6 -ICPU=3 -NO MAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100

-LONGLOG -DRV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5

-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

- 1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq.*
- 2: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq.*
- 3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq.*
- 4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq.*
- 5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
- 6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq.*
- 7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
- 8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq.*
- 9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq.*
- 10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq.*
- 11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq.*
- 12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
- 13: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
- 14: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq.*
- 15: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq.*
- 16: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq.*
- 17: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq.*
- 18: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq.*
- 19: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
------------	-------	-------------	--------	----	-------------

ALIGNMENTS

RESULT 1

US-10-081-775-1

Sequence 1, Application US/10081775

Publication No. US20030060409A1

GENERAL INFORMATION:

APPLICANT: Bristol-Myers Squibb Company

TITLE OF INVENTION: POLYNUCLEOTIDE ENCODING A NOVEL HUMAN G-PROTEIN COUPLED RECEPTOR, HGPBMY25, EXPRESSED HIGHLY IN IMMUNE-RELATED TISSUES

FILE REFERENCE: D0126 NP

CURRENT APPLICATION NUMBER: US/10/081,775

PRIOR FILING DATE: 2002-02-21

PRIOR FILING DATE: 2001-02-21

PRIOR APPLICATION NUMBER: US 60/270,134

PRIOR FILING DATE: 2001-03-27

NUMBER OF SEQ ID NOS: 69

SOFTWARE: PatentIn version 3.0

SEQ ID NO 1

LENGTH: 1567

TYPE: DNA

ORGANISM: homo sapiens

FEATURE:

NAME/KEY: CDS

1	1718	100.0	1567	15	US-10-081-775-1	Sequence 1, Appli
2	1691	98.4	990	9	US-09-886-055-262	Sequence 262, App
3	1691	98.4	990	10	US-09-804-291-262	Sequence 262, App
4	1691	98.4	990	13	US-10-343-650A-353	Sequence 353, App
5	1691	98.4	998	15	US-10-025-806-35	Sequence 35, Appl
6	1691	98.4	998	15	US-10-025-806-37	Sequence 37, Appl
7	1691	98.4	1366	15	US-10-017-161-923	Sequence 923, App
8	1691	98.4	1390	16	US-10-292-798-797	Sequence 797, App
9	1649	96.0	966	17	US-10-297-021-42	Sequence 42, Appl
10	978.5	57.0	1300	16	US-10-292-798-833	Sequence 833, App
11	951.5	55.4	1242	15	US-10-017-161-959	Sequence 959, App
12	931.5	54.2	990	13	US-10-343-650A-315	Sequence 315, App
13	931.5	54.2	990	16	US-10-387-629-157	Sequence 157, App
14	931.5	54.2	1114	17	US-10-467-252-93	Sequence 93, Appl
15	931.5	54.2	1390	15	US-10-017-161-945	Sequence 945, App
16	931.5	54.2	1390	16	US-10-292-798-819	Sequence 819, App
17	920	53.6	1076	13	US-10-182-822A-23	Sequence 23, Appl
18	907	52.8	945	15	US-10-044-643-19	Sequence 19, Appl
19	904	52.6	936	9	US-09-886-055-114	Sequence 114, App
20	904	52.6	936	10	US-09-804-291-114	Sequence 114, App
21	904	52.6	936	13	US-10-343-650A-213	Sequence 213, App
22	904	52.6	985	15	US-10-044-643-15	Sequence 15, Appl
23	904	52.6	1001	15	US-10-220-382-31	Sequence 31, Appl
24	904	52.6	1258	15	US-10-017-161-909	Sequence 909, App
25	904	52.6	1336	16	US-10-292-798-783	Sequence 783, App
26	903	52.6	947	15	US-10-044-643-17	Sequence 17, Appl
27	901.5	52.5	1013	15	US-10-025-806-25	Sequence 25, Appl
28	899.5	52.4	1360	15	US-10-017-161-947	Sequence 947, App
29	899.5	52.4	1360	16	US-10-292-798-821	Sequence 821, App
30	889.5	51.8	957	9	US-09-886-055-102	Sequence 102, App
31	889.5	51.8	957	10	US-09-804-291-102	Sequence 102, App
32	889.5	51.8	957	13	US-10-343-650A-197	Sequence 197, App
33	889.5	51.8	990	15	US-10-025-806-1	Sequence 1, Appl
34	889.5	51.8	1345	15	US-10-017-161-877	Sequence 877, App
35	888.5	51.7	980	15	US-10-025-806-3	Sequence 3, Appl
36	886.5	51.6	982	11	US-09-844-861A-21	Sequence 21, Appl
37	886.5	51.6	1592	17	US-10-466-720-29	Sequence 29, Appl
38	885.5	51.5	966	9	US-09-886-055-260	Sequence 260, App
39	885.5	51.5	966	13	US-09-804-291-260	Sequence 189, App
40	885.5	51.5	966	15	US-10-343-650A-189	Sequence 189, App
41	885.5	51.5	971	16	US-10-024-212-91	Sequence 91, Appl
42	885.5	51.5	1246	15	US-10-017-161-925	Sequence 925, App
43	885.5	51.5	1366	16	US-10-292-798-799	Sequence 799, App
44	884.5	51.5	945	13	US-10-343-650A-159	Sequence 159, App
45	884.5	51.5	1151	11	US-09-844-861A-19	Sequence 19, Appl

Db 481 CGAGGCTCCCTGCTCATGCTCCCATGCTCTCTTTATTGGGCGTTGACATTCGCCAA 540
Qy 181 SerHisValIleuHisThrTyrcysGluHisMetAlaValVallysLeuAlaCysGly 200
Db 541 ACCATGTGATCTCTACACAGTACTGTGAGCATGCTGTGGTGAAGCTGGCTGTGGA 600
Qy 201 AspThrArgProAsnArgValTyrcysGluHisMetAlaValVallysLeuAlaCysGly 220
Db 601 GACACAGGCTTAAACGCTGTATGGCTGACAGCTGCACCTGTGTGTCATTGGGGTGC 660
Qy 221 LeuPheCysIleGlyLeuSerTyrcysGluHisMetAlaValVallysLeuAlaCysGly 240
Db 661 TTGTTTTCATGCTCTCTCTATGCTTAAAGTGCACAGCTGCTCTGCTCTCTATCC 720
Qy 241 HisGluAlaArgSerIlysAlaLeuGlyThrCysGlySerHisValCysVallIleLeu 260
Db 721 CATGAAGCTCGGTGCAAGGCTTAGGACCTGTGGTTCCCATGCTGTGTCTATCCATC 780
Qy 261 SerTyrcysAlaLeuPheSerPhePheThrHisArgPheGlyHisValProVal 280
Db 781 TCTTATACACAGGCTCTCTCTCTTTTACACACGCTTTGGCCATCAGCTTCCAGTC 840
Qy 281 HisIleHisIleLeuAlaAsnValTyrcysGluHisMetAlaValVallysLeuAlaCysGly 300
Db 841 CATATTTCATCTCTTTGGCCAAATGTTATCTGCTTTTGGCCACCTGCTTAACTCTGTG 900
Qy 301 ValTyrcysValTyrcysGlnIleArgLysArgValValArgValPheGlnSerGly 320
Db 901 GTATATGAGTTAAACCAACAGATCCGTAAAGAGTTGTCAAGGCTGTTTCAAAGTGG 960
Qy 321 GlnGlyMetGlyIleLysAlaSerGlu 329
Db 961 CAGGAATGGGCATCAAGGCATCTGAG 987

RESULT 3

US-09-804-291-262
; Sequence 262, Application US/09804291
; Publication No. US20030088059A1
; GENERAL INFORMATION:
; APPLICANT: ZOZULA, SERGEY
; TITLE OF INVENTION: HUMAN OLFACTORY RECEPTORS AND GENES ENCODING SAME
; FILE REFERENCE: P 0278005
; CURRENT APPLICATION NUMBER: US/09/804,291
; PRIOR FILING DATE: 2001-03-13
; PRIOR APPLICATION NUMBER: 60/188,914
; PRIOR FILING DATE: 2000-03-13
; PRIOR APPLICATION NUMBER: 60/192,033
; PRIOR FILING DATE: 2000-03-24
; PRIOR APPLICATION NUMBER: 60/198,474
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/199,335
; PRIOR FILING DATE: 2000-04-24
; PRIOR APPLICATION NUMBER: 60/207,702
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/213,849
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: 60/226,534
; PRIOR FILING DATE: 2000-08-16
; PRIOR APPLICATION NUMBER: 60/230,732
; PRIOR FILING DATE: 2000-09-07
; PRIOR APPLICATION NUMBER: 60/266,862
; PRIOR FILING DATE: 2001-02-07
; NUMBER OF SEQ ID NOS: 529
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 262
; LENGTH: 990
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-804-291-262

Alignment Scores:

Pred. No.: 1.21e-192 Length: 990

Score: 1691.00 Matches: 325
Percent Similarity: 99.09% Conservative: 1
Best Local Similarity: 98.78% Mismatches: 3
Query Match: 10 Indels: 0
DB: 10 Gaps: 0

US-10-081-775-2 (1-329) x US-09-804-291-262 (1-990)

Qy 1 MetSerSerThrLeuGlyHisAsnMetGluSerProHisHisThrAspValAspProSer 20
Db 1 ATGTCCAGCACTCTTGGCCACAACTGAATCTCTAATCACACATGATGTTGACCCCTTCT 60
Qy 21 ValPhePheLeuGlyIleProGlyLeuGluGlnPheHisLeuTrpLeuSerLeuPro 40
Db 61 GTCCTTCTCTCTGCGCATCCAGGTCTGGAACAATTTCAATTTGGGCTCTCACTCCCT 120
Qy 41 ValCysGlyLeuGlyThrAlaThrIleValGlyAsnIleThrIleLeuValValAla 60
Db 121 GGTGTGGCTTAGGCACAGCCCAATTTGGGCAATATAAATAATTTCTGTTGTTGTC 180
Qy 61 ThrGluProValLeuHisLysProValTyrcysGluHisMetLeuSerThrIleAsp 80
Db 181 ACTGAACCACTTTCACACAGCTGTGTACCTTTTCTGTGATGCTCTCAACCACTCGAC 240
Qy 81 LeuAlaAlaSerValSerThrValProLysLeuLeuAlaIlePheTrpCysGlyAlaGly 100
Db 241 TTGGCTGCTCTGTCTCCACAGTTCACCAAGCTACTGGCTATCTTCTGTGTGGAGCCGA 300
Qy 101 HisIleSerAlaSerAlaCysLeuAlaHisMetPhePheIleHisAlaPheCysMetMet 120
Db 301 CATATATCTGCTCTGCTGCTGCTGSCACAGATGTTCTTCATTTTCATGCTCTGCAATG 360
Qy 121 GluSerThrValLeuAlaMetAlaPheAspArgTyrcysGluHisMetLeuSerProLeu 140
Db 361 GAGTCCACTGTGCTACTGGCCATGCTTGTGATCGTACGTGGCCATCTGCCACCACTTC 420
Qy 141 ArgTyrcysAlaThrIleLeuThrAspThrIleAlaHisIleGlyValAlaValVal 160
Db 421 CGTATGTCACAACTCTCACTGACACCATCATTTGCCACATAGGGGTGGAGCTGTAGTG 480
Qy 161 ArgGlySerLeuMetLeuProCysProPhePheIleGlyArgLeuAsnPheCysGln 180
Db 481 CGAGCTCCCTGCTCATGCTCCCATGCTCTCTTATTTGGGCGTTTGAATCTTCGCA 540
Qy 181 SerHisValIleLeuHisThrTyrcysGluHisMetAlaValVallysLeuAlaCysGly 200
Db 541 AGCCATGTGATCTTACACAGTACTGTGAGCACATGGCTGTGGTGAAGCTGGCTGTGGA 600
Qy 201 AspThrArgProAsnArgValTyrcysGluHisMetAlaValVallysLeuAlaCysGly 220
Db 601 GACACAGGCTTAAACGCTGTATGGCTGACAGCTGCTGTTGGTCAATTTGGGCTTGC 660
Qy 221 LeuPheCysIleGlyLeuSerTyrcysGluHisMetAlaValVallysLeuAlaCysGly 240
Db 661 TTGTTTTCATGCTCTCTCTATGCTTAAAGTGCACAGCTGCTCTGCTCTCTATCC 720
Qy 241 HisGluAlaArgSerIlysAlaLeuGlyThrCysGlySerHisValCysVallIleLeu 260
Db 721 CATGAAGCTCGGTGCAAGGCTTAGGACCTGTGGTTCCCATGCTGTGTCTATCCATC 780
Qy 261 SerTyrcysAlaLeuPheSerPhePheThrHisArgPheGlyHisValProVal 280
Db 781 TCTTATACACAGGCTCTCTCTCTTTTACACACGCTTTGGCCATCAGCTTCCAGTC 840
Qy 281 HisIleHisIleLeuAlaAsnValTyrcysGluHisMetLeuSerProAlaLeuAsnProVal 300
Db 841 CATATTTCATCTCTTTGGCCAAATGTTATCTGCTTTTGGCCACCTGCTTAACTCTGTG 900
Qy 301 ValTyrcysValTyrcysGlnIleArgLysArgValValArgValPheGlnSerGly 320
Db 901 GTATATGAGTTAAACCAACAGATCCGTAAAGAGTTGTCAAGGCTGTTTCAAAGTGG 960
Qy 321 GlnGlyMetGlyIleLysAlaSerGlu 329

```
Db 961 CAGGGAATGGGCATCAAGGCATCTGAG 987
|||||
RESULT 4
US-10-343-650A-353
; Sequence 353, Application US/10343650A
; Publication No. US20040067499A1
; GENERAL INFORMATION:
; APPLICANT: HAGA, TATSUYA
; TITLE OF INVENTION: NOVEL G-PROTEIN COUPLED RECEPTOR
; FILE REFERENCE: 31671-186347
; CURRENT APPLICATION NUMBER: US/10/343,650A
; CURRENT FILING DATE: 2003-07-21
; PRIOR APPLICATION NUMBER: JP 2000/237818
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: JP 2001/34434
; PRIOR FILING DATE: 2001-02-13
; NUMBER OF SEQ ID NOS: 694
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 353
; LENGTH: 990
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(990)
US-10-343-650A-353

Alignment Scores:
Pred. No.: 1,218-192 Length: 990
Score: 1691.00 Matches: 325
Percent Similarity: 99.08% Conservative: 1
Best Local Similarity: 98.78% Mismatches: 3
Query Match: 98.43% Indels: 0
DB: 13 Gaps: 0

US-10-081-775-2 (1-329) x US-10-343-650A-353 (1-990)
```

```
QY 1 MetSerSerThrLeuGlyHisAsnMetGluSerProHisHisThrAspValAspProSer 20
|||
Db 1 ATGTCAGCAGCTCTTGGGCACACATGGAATCTCTTAATCACACATGATGTTGACCCCTCT 60
|||
QY 21 ValPhePheLeuLeuGlyProGlyLeuGluGlnPheHisLeuTrpLeuSerLeuPro 40
|||
Db 61 GTCCTCTCTCTGCGCATCCAGCTCTGGAAACATTTTCATTTGGGCTCTCATCCCT 120
|||
QY 41 ValCysGlyLeuGlyThrAlaThrIleValGlyAsnIleThrIleLeuValValAla 60
|||
Db 121 GTGTGTGGCTTAGGCACACCCACAAATTTGGGCAATATAACTATTCTGGTGTGTTGCC 180
|||
QY 61 ThrGluProValLeuHisIleValProValThrLeuPheLeuCysMetLeuSerThrIleAsp 80
|||
Db 181 ACTGAACCAAGCTCTGCACAAAGCTGTGTACCTTTTCTGTGCATGCTCTCAACCATCGAC 240
|||
QY 81 LeuAlaAlaSerValSerThrValProLysLeuAlaIlePheTrpCysGlyAlaGly 100
|||
Db 241 TTGGCTGCTCTGTCTCCACAGTTCCCAAGCTACTGCTATCTTCTGGTGTGGAGCCGA 300
|||
QY 101 HisIleSerAlaSerAlaCysLeuAlaHisMetPheIleHisAlaPheCysMetMet 120
|||
Db 301 CATATATCTGCCTCTGCTGCTGGCGGCACAGATGTTCTTCATTCATGCTTCTGCATGATG 360
|||
QY 121 GluSerThrValLeuLeuAlaMetAlaPheAspArgTrpValAlaIleCysHisProLeu 140
|||
Db 361 GAGTCCACTGTGTACTGGCCATGGGCTTTGATCGCTACGTGGCCATCTGCCACCCATC 420
|||
QY 141 ArgTrpAlaThrIleLeuThrAspThrIleAlaHisIleGlyValAlaValVal 160
|||
Db 421 CGTATGCCACATCTCTCATGACATCATTTGCCACATAGGGTGGCAGGTGTAGTG 480
|||
QY 161 ArgGlySerLeuLeuMetLeuProCysProPheIleGlyArgLeuAsnPheCysGln 180
|||
Db 481 CGAGGCTCCCTGCTCATGCTCCCATGTCCTTCTTATTGGGCGTTTGAACCTCTGCCAA 540
|||
```

```
QY 181 SerHisValIleLeuHisThrTyrCysGluHisMetAlaValValLysLeuAlaCysGly 200
|||
Db 541 AGCCATGTGATCTTACACACAGTACTGTGAGCACATGGCTGTGGTGAAGCTGGCTGTGGA 600
|||
QY 201 AspThrArgProAsnArgValTyrGlyLeuThrAlaAlaLeuLeuValIleGlyValAsp 220
|||
Db 601 GACACACGGCCCTAACCGGTGTATGGGCTGACAGCTGCACCTGTTGGTCAATTGGGGTTGAC 660
|||
QY 221 LeuPheCysIleGlyLeuSerTyrAlaLeuIleAlaGlnAlaValLeuArgLeuSerSer 240
|||
Db 661 TTGTTTTGCATTTGGTCTCTCTATGCCCTAAGTGACAAAGCTGCTTCCCTTCATCC 720
|||
QY 241 HisGluAlaArgSerLysAlaLeuGlyThrCysGlySerHisValCysValIleLeuIle 260
|||
Db 721 CATGAGCTCGTCCAGGCCCTTAGGACCTGTGGTCCCATGCTGTGTGTCATCTCATC 780
|||
QY 261 SerTyrThrProAlaLeuPheSerPhePheThrHisArgPheGlyHisHisValProVal 280
|||
Db 781 TCTTATACACACAGCCCTCTCTCTCTTTTACACACCCGCTTTGGCCATCAGCTTCCAGTC 840
|||
QY 281 HisIleHisIleLeuLeuAlaAsnValTyrLeuLeuLeuProAlaLeuAsnProVal 300
|||
Db 841 CATATTACATTTCTTTGGCCAAATGTTATCTGCTTTTGGCCACCTGCTCTTAATCTGTG 900
|||
QY 301 ValTyrGlyValLysThrLysGlnIleArgLysArgValValArgValPheGlnSerGly 320
|||
Db 901 GTATATGAGATTAGACCAACAGATCCGTAAGAGAGTTGTCAGGGGTGTTTCAAGTGGG 960
|||
QY 321 GlnGlyMetGlyIleLysAlaSerGlu 329
|||
Db 961 CAGGGAATGGGCATCAAGGCATCTGAG 987
|||
RESULT 5
US-10-025-806-35
; Sequence 35, Application US/10025806
; Publication No. US20030198955A1
; GENERAL INFORMATION:
; APPLICANT: Li, Li
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Ballinger, Robert
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Colman, Steven
; APPLICANT: Spytek, Kimberly
; APPLICANT: Casman, Stacie
; APPLICANT: Edinger, Shlomit
; APPLICANT: Gerlach, Valerie
; APPLICANT: Sciore, Paul
; APPLICANT: Smithson, Glennnda
; APPLICANT: Peyman, John
; APPLICANT: MacDougall, John
; APPLICANT: Stone, David
; APPLICANT: Vernet, Corine
; APPLICANT: Shenoy, Suresh
; APPLICANT: Gunther, Erik
; APPLICANT: Millet, Isabelle
; APPLICANT: Tchernev, Velizar
; APPLICANT: Anderson, David
; APPLICANT: Gusev, Vladimir
; APPLICANT: Malyankar, Uriel
; APPLICANT: Zhong, Haihong
; APPLICANT: Ellerman, Karen
; APPLICANT: Wolenc, Adam
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 21402-224 AB
; CURRENT APPLICATION NUMBER: US/10/025,806
; CURRENT FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 60/256,635
; PRIOR FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: 60/259,743
; PRIOR FILING DATE: 2001-01-04
; PRIOR APPLICATION NUMBER: 60/299,327
; PRIOR FILING DATE: 2001-06-19
```

; PRIOR APPLICATION NUMBER: 60/261,498
 ; PRIOR FILING DATE: 2001-01-12
 ; PRIOR APPLICATION NUMBER: 60/263,689
 ; PRIOR FILING DATE: 2001-01-24
 ; PRIOR APPLICATION NUMBER: 60/276,464
 ; PRIOR FILING DATE: 2001-02-08
 ; PRIOR APPLICATION NUMBER: 60/271,021
 ; PRIOR FILING DATE: 2001-02-22
 ; PRIOR APPLICATION NUMBER: 60/275,946
 ; PRIOR FILING DATE: 2001-03-14
 ; PRIOR APPLICATION NUMBER: 60/278,150
 ; PRIOR FILING DATE: 2001-03-23
 ; PRIOR APPLICATION NUMBER: 60/285,718
 ; PRIOR FILING DATE: 2001-04-23
 ; PRIOR APPLICATION NUMBER: 60/312,902
 ; PRIOR FILING DATE: 2001-08-16
 ; PRIOR APPLICATION NUMBER: 60/257,876
 ; PRIOR FILING DATE: 2000-12-21
 ; PRIOR APPLICATION NUMBER: 60/260,718
 ; PRIOR FILING DATE: 2001-01-10
 ; PRIOR APPLICATION NUMBER: 60/284,591
 ; PRIOR FILING DATE: 2001-04-18
 ; NUMBER OF SEQ ID NOS: 352
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 35
 ; LENGTH: 998
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (2)..(998)
 US-10-025-806-35

Alignment Scores:
 Pred. No.: 1,22e-192 Length: 998
 Score: 1691.00 Matches: 325
 Percent Similarity: 99.09% Conservative: 1
 Best Local Similarity: 98.78% Mismatches: 3
 Query Match: 98.43% Indels: 0
 DB: 15 Gaps: 0

US-10-081-775-2 (1-329) x US-10-025-806-35 (1-998)

Qy 1 MetSerThrLeuGlyHisAsnMetGluSerProHisHisThrAspValaspProSer 20
 Db 2 ATGTCAGACATCTTGGCCACAAATGGAATCTCTTAATCACATGATGTTGACCTTCT 61
 Qy 21 ValPhePheLeuLeuGlyIleProGlyLeuGluGlnPheHisLeuTrpLeuSerLeuPro 40
 Db 62 GTCTTCTTCTCTCTGGGCATCCAGGCTCGAACAAATTTCAATTTGGCTCTCACTCCCT 121
 Qy 41 ValCysGlyLeuGlyThrAlaThrIleValGlyAsnIleThrIleLeuValValAla 60
 Db 122 GTGTGGCTTAGGCACAGCCACAAATTTGGGCAATATAACTATCTGGTTGTTGGCC 181
 Qy 61 ThrGluProValLeuHisLysProValTrpLeuPheLeuCysMetLeuSerThrIleAsp 80
 Db 182 ACTGAACAGCTGTGCACAGCTGTGTACCTTTTCTGTGCATGCTCTCAACCATGCAC 241
 Qy 81 LeuAlaAlaSerValSerThrValProLysLeuLeuAlaIlePheTrpCysGlyAlaGly 100
 Db 242 TTGGCTGCCTCTGTCTCCACAGTTCACAGCTACTGGCTATCTTCTGGTGTGGAGCCGA 301
 Qy 101 HistSerAlaSerAlaCysLeuAlaHisMetPhePheIleHisAlaPheCysMetMet 120
 Db 302 CATATATCTGCCTCTGCCTGCCAGAGATGTCTTCTATTCATGCTTCTGCATGATG 361
 Qy 121 GluSerThrValLeuLeuAlaMetAlaPheAspArgTrpValAlaIleCysHisProLeu 140
 Db 362 GAGTCCACTGTGCTACTGGCCATGGCCCTTTGATCGCTACGTGGCCATCTGCCACCCACTC 421
 Qy 141 ArgTrpAlaThrIleLeuThrAspThrIleIleAlaHisIleGlyValAlaValVal 160

Db 422 CGCTATGCCACAATCTCTACTGACACCATCATATTGCCCATAGGGGTGGCAGCTGTAGTG 481
 Qy 161 ArgGlySerLeuLeuMetLeuProCysProPhePheIleGlyArgLeuAsnPheCysGln 180
 Db 482 CGAGGCTCCTCTCATCTCCCATGTCCTTCTTATTTGGGCGTTTGACCTTCTGCCAA 541
 Qy 181 SerHisValIleLeuHisThrTrpCysGluHisMetAlaValValLysLeuAlaCysGly 200
 Db 542 AGCCATGTGATCTACACAGCTACTGTGAGCACATGGCTGTGGTGAAGCTGGCTGTGGA 601
 Qy 201 AspThrArgProAsnArgValTrpGlyLeuThrAlaAlaLeuValIleGlyValAsp 220
 Db 602 GACACAGCCCTAACCGTGTGTATGGCTGTGACAGCTGCACCTGTGTGCTATTGGGGTTGAC 661
 Qy 221 LeuPheCysIleGlyLeuSerTrpAlaLeuIleAlaGlnAlaValLeuArgLeuSerSer 240
 Db 662 TTGTTTTCATTTGGTCTCTCTCTATGCCCTAAGTCACAGCTGTCTTTCGCCCTCTCATCC 721
 Qy 241 HisGluAlaArgSerLysAlaLeuGlyThrCysGlySerHisValCysValIleLeuIle 260
 Db 722 CATGAAGCTCGTCCAGGCCCTTAGGGACCTGTGGTTCCTCATGTCTGTGTATCTCATC 781
 Qy 261 SerTrpThrProAlaLeuPheSerPhePheThrHisArgPheGlyHisValProVal 280
 Db 782 TCTTATACACAGCCCTCTTCTCTTTTACACACCCGCTTTGGCCATCACGTTCCAGTC 841
 Qy 281 HistHisIleLeuLeuAlaAsnValTrpLeuLeuProAlaLeuAsnProVal 300
 Db 842 CATATTACATTTCTTTGGCCAAATGTTTATCTGTCTTTGGCCACCTGCTCTTAATCTCTG 901
 Qy 301 ValTrpGlyValLysThrLysGlnIleArgLysArgValValArgValPheGlnSerGly 320
 Db 902 GTATATGGAGTTAAGACCAAAACAGATCCGTAAGAGAGTTGTCAAGGGTGTTCCTTCAAGTGGG 961
 Qy 321 GlnGlyMetGlyIleLysAlaSerGlu 329
 Db 962 CAGGGAATGGGCATCAGGCATCTGAG 988

RESULT 6

US-10-025-806-37
 ; Sequence 37, Application US/10025806
 ; Publication No. US20030198955A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Li, Li
 ; APPLICANT: Padigaru, Muralidhara
 ; APPLICANT: Ballinger, Robert
 ; APPLICANT: Kekuda, Ramesh
 ; APPLICANT: Colman, Steven
 ; APPLICANT: Spytek, Kimberly
 ; APPLICANT: Casman, Stacie
 ; APPLICANT: Edinger, Shlomit
 ; APPLICANT: Gerlach, Valerie
 ; APPLICANT: Sciore, Paul
 ; APPLICANT: Smithson, Glenda
 ; APPLICANT: Peyman, John
 ; APPLICANT: MacDougall, John
 ; APPLICANT: Stone, David
 ; APPLICANT: Vernet, Corine
 ; APPLICANT: Shenoy, Suresh
 ; APPLICANT: Gunther, Erik
 ; APPLICANT: Millet, Isabelle
 ; APPLICANT: Tchernev, Velizar
 ; APPLICANT: Anderson, David
 ; APPLICANT: Gusev, Vladimir
 ; APPLICANT: Malyankar, Uriel
 ; APPLICANT: Zhong, Haihong
 ; APPLICANT: Ellerman, Karen
 ; APPLICANT: Wolenc, Adam
 ; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
 ; FILE REFERENCE: 21402-224 AB
 ; CURRENT APPLICATION NUMBER: US/10/025,806
 ; CURRENT FILING DATE: 2001-12-19
 ; PRIOR APPLICATION NUMBER: 60/256,635

```

; PRIOR FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: 60/259,743
; PRIOR FILING DATE: 2001-01-04
; PRIOR APPLICATION NUMBER: 60/299,327
; PRIOR FILING DATE: 2001-06-19
; PRIOR APPLICATION NUMBER: 60/261,498
; PRIOR FILING DATE: 2001-01-12
; PRIOR APPLICATION NUMBER: 60/263,689
; PRIOR FILING DATE: 2001-01-24
; PRIOR APPLICATION NUMBER: 60/276,464
; PRIOR FILING DATE: 2001-02-08
; PRIOR APPLICATION NUMBER: 60/271,021
; PRIOR FILING DATE: 2001-02-22
; PRIOR APPLICATION NUMBER: 60/275,946
; PRIOR FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: 60/278,150
; PRIOR FILING DATE: 2001-03-23
; PRIOR APPLICATION NUMBER: 60/285,718
; PRIOR FILING DATE: 2001-04-23
; PRIOR APPLICATION NUMBER: 60/312,902
; PRIOR FILING DATE: 2001-08-16
; PRIOR APPLICATION NUMBER: 60/257,876
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/260,718
; PRIOR FILING DATE: 2001-01-10
; PRIOR APPLICATION NUMBER: 60/284,591
; PRIOR FILING DATE: 2001-04-18
; NUMBER OF SEQ ID NOS: 352
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 37
; LENGTH: 998
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (2)..(988)
US-10-025-806-37

Alignment Scores:
Pred. No.: 1.22e-192 Length: 998
Score: 1691.00 Matches: 325
Percent Similarity: 99.09% Conservative: 1
Best Local Similarity: 98.78% Mismatches: 3
Query Match: 98.43% Indels: 0
DB: 15 Gaps: 0

US-10-081-775-2 (1-329) x US-10-025-806-37 (1-998)

QY 1 MetSerThrLeuGlyHisAlaMetGluSerProHisHisThrAspValAspProSer 20
Db 2 ATGTCAGCACCTTTGGCCACACATGGAAATCTCCTAATCACACTGATGTTGACCCCTTCT 61
QY 21 ValPhePheLeuLeuGlyProGlyLeuGluGlnPheHisHisLeuTrpLeuSerLeuPro 40
Db 62 GTCTTCTTCTCTGGCATCCAGGCTCGGAACAATTTCAITTTGGGTCTCACTCCCT 121
QY 41 ValCysGlyLeuGlyThrAlaThrLeuValGlyAsnLeuThrLeuValValAla 60
Db 122 GTGTGTGGCTTAGGCACACCAATTTGGGCAATATAACTATCTGTGTGTGTGTGCC 181
QY 61 ThrGluProValLeuHisLysProValTyrLeuPheLeuCysMetLeuSerThrLeuAsp 80
Db 182 ACTGAACCAAGCTTGGCAGAGCTGTGTACCTTTTCTGTGCATGCTCTCAACCATCGAC 241
QY 81 LeuAlaAlaSerValSerThrValProLysLeuLeuAlaLeuPheTrpCysGlyAlaGly 100
Db 242 TTGGTCTGCTCTGCTCCACAGTCCCAAGCTACTGGCTATCTTCTGTGTGTGGAGCCGA 301
QY 101 HisLeuSerAlaSerAlaCysLeuAlaHisMetPhePheLeuHisAlaPheCysMetMet 120
Db 302 CATATATCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 361
QY 121 GluSerThrValLeuLeuAlaMetAlaPheAspArgTyrValAlaLeuCysHisProLeu 140

```

```

Db 362 GAGTCCACTGTGCTACTGGCCATGGCTTGTGATCGTACGTGGCCATCTGCCACCCACTC 421
QY 141 ArgTyrAlaThrLeuLeuThrAspThrLeuLeuAlaHisLeuGlyValAlaAlaValVal 160
Db 422 CGCTATGCCAATCTCTCACTGACACCATCATGCCACATAGGGGTGGAGCTGTAGTG 481
QY 161 ArgGlySerLeuLeuMetLeuProCysProPhePheLeuGlyArgLeuAsnPheCysGln 180
Db 482 CGAGGCTCCCTGCTCATGCTCCCATGTCCTTCTTATTGGCGCTTTGAACCTCTGCCAA 541
QY 181 SerHisValLeuLeuHisThrTyrCysGluHisMetAlaValValLysLeuAlaCysGly 200
Db 542 AGCCATGTGATCTTACACACGCTACTGTGAGCACATGGCTGTGGTGAAGCTGGCTGTGGA 601
QY 201 AspThrArgProAsnArgValTyrGlyLeuThrAlaLeuLeuValLeuGlyValAsp 220
Db 602 GACACACAGGCTTAACCGCTGTATGGGCTGACAGCTGCTGTGTGTGTGTGTGTGTGTGTG 661
QY 221 LeuPheCysLeuGlyLeuSerTyrAlaLeuLeuAlaGlnAlaValLeuArgLeuSerSer 240
Db 662 TTGTTTTCATGTTGCTCTCTCTATGCCCTAAGTGCACAAAGCTGCTCTGCTCTCTCATCC 721
QY 241 HisGluAlaArgSerLysAlaLeuGlyThrCysGlySerHisValCysValLeuLeuLeu 260
Db 722 CATGAAGCTCGGTCCAGGCTTAGGACCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 781
QY 261 SerTyrThrProAlaLeuPheSerPhePheThrHisArgPheGlyHisValProVal 280
Db 782 TCTTATACACAGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 841
QY 281 HisLeuHisLeuLeuAlaAsnValTyrLeuLeuLeuLeuProAlaLeuAsnProVal 300
Db 842 CATATTCATCTCTTTGGCCAAATGTTATCTGCTTTTGGCCACCTGCTCTTAATCTGTG 901
QY 301 ValTyrGlyValLysThrLysGlnLeuArgLysArgValValArgValPheGlnSerGly 320
Db 902 GTATATGGAGTTAAGACCAACACAGATCGTAAAGAGTTGTGAGGTGTGTTTCAAAAGTGGG 961

RESULT 7
US-10-017-161-923
; Sequence 923, Application US/10017161
; Publication NO. US20030143668A1
; GENERAL INFORMATION:
; APPLICANT: SUWA, MAKIKO
; APPLICANT: ASAI, KIYOSHI
; APPLICANT: AKIYAMA, YUTAKA
; APPLICANT: ABURATANI, HIROYUKI
; TITLE OF INVENTION: NOVEL G PROTEIN-COUPLED RECEPTORS
; FILE REFERENCE: 084335/0152
; CURRENT APPLICATION NUMBER: US/10/017,161
; CURRENT FILING DATE: 2002-12-18
; PRIOR APPLICATION NUMBER: JP 2001/246789
; PRIOR FILING DATE: 2001-06-18
; NUMBER OF SEQ ID NOS: 2430
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 923
; LENGTH: 1366
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: source
; LOCATION: (1)..(1366)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (201)..(1166)
US-10-017-161-923

```

Alignment Scores:

Pred. No.: 1.99e-192 Length: 1366
Score: 1691.00 Matches: 325
Percent Similarity: 99.09% Conservatives: 1
Best Local Similarity: 98.78% Mismatches: 3
Query Match: 98.43% Indels: 0
DB: 15 Gaps: 0

US-10-081-775-2 (1-329) x US-10-017-161-923 (1-1366)

QY 1 MetSerSerThrLeuGlyHisAsnMetGluSerProHisHisThrAspValAspProSer 20
DB 177 ATGTCAGACACTCTTGGCCACACATGGAATCTCTTAATCACACTGATGTGACCTTCT 236
QY 21 ValPhePheLeuLeuGlyLeuProGlyLeuGluGlnPheHisLeuTrpLeuSerLeuPro 40
DB 237 GTCCTTCTCTCTGGGCATCCAGGCTGGGAACAATTTCAATTTGGGCTCTCACTCCCT 296
QY 41 ValCysGlyLeuGlyThrAlaThrIleValGlyAsnIleThrIleLeuValValAla 60
DB 297 GTGTGTGGCTTAGGCACAGCCCAATTTGGGCAATATAACTATCTGTTGTTGCC 356
QY 61 ThrGluProValLeuHisLysProValTyrLeuPheLeuCysMetLeuSerThrIleAsp 80
DB 357 ACTGAACACAGCTTTGCACAAAGCTGTGTACCTTTTCTGTGATGCTCTCAACCATCGAC 416
QY 81 LeuAlaAlaSerValSerThrValProLysLeuLeuAlaIlePheTrpCysGlyAlaGly 100
DB 417 TTGGCTGCTCTGCTCCACAGTTCCCAAGCTACTGGCTATCTTCTGGTGTGGAGCCGGA 476
QY 101 HisIleSerAlaSerAlaCysLeuAlaHisMetPhePheIleHisAlaPheCysMetMet 120
DB 477 CATATATCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 536
QY 121 GluSerThrValLeuLeuAlaMetAlaPheAspArgTyrValAlaIleCysHisProLeu 140
DB 537 GAGTCCACTGTGCTACTGGCCATGGCCCTTGTATCGCTAGTGGCCATCTGCCACCATCT 596
QY 141 ArgTyrAlaThrIleLeuThrAspThrIleIleAlaHisIleGlyValAlaAlaValVal 160
DB 597 CGCTATGACCAATCTCTACACACATCTGTCAGCAGTCTGCTGCTGCTGCTGCTGCTG 656
QY 161 ArgGlySerLeuLeuMetLeuProCysProPhePheIleGlyArgLeuAsnPheCysGln 180
DB 657 CGAGGCTCCCTGCTCATGCTCCCATGCTCCCTTCTTATTTGGGCGTTGAATCTTGCCAA 716
QY 181 SerHisValIleLeuHisThrTyrCysGluHisMetAlaValValLysLeuAlaCysGly 200
DB 717 AGCCATGTGATCTTACACACAGTACTGTGAGCAGCATGGCTGTGTGAAGCTGGCTGTGA 776
QY 201 AspThrArgProAsnArgValTyrGlyLeuThrAlaAlaLeuLeuValIleGlyValAsp 220
DB 777 GACACACAGGCTAACCGTGTGTATGGGCTGACAGCTGCTGCTGCTGCTGCTGCTGCTG 836
QY 221 LeuPheCysIleGlyLeuSerTyrAlaLeuIleAlaGlnAlaValLeuArgLeuSerSer 240
DB 837 TTGTTTGTGATTTGCTCTCTCTATGCTTAAAGTGCACAGCTGCTGCTGCTGCTGCTG 896
QY 241 HisGluAlaArgSerLysAlaLeuGlyThrCysGlySerHisValCysValIleLeuIle 260
DB 897 CATGAAGCTGGTCCAAAGGCTTAGGACCTGTGGTTCCTGCTGCTGCTGCTGCTGCTGCT 956
QY 261 SerTyrThrProAlaLeuPheSerPhePheThrHisArgPheGlyHisHisValProVal 280
DB 957 TCTTATACACAGGCTCTTCTCTCTTTTATACACACGCTTTGGCCATCAGCTTCCAGTC 1016
QY 281 HisIleHisIleLeuLeuAlaAsnValTyrLeuLeuLeuProProAlaLeuAsnProVal 300
DB 1017 CATATTACATCTTTTGGCCAAATGTTATCTGCTTTTGGACCTGCTCTTAATCTCTGTG 1076
QY 301 ValTyrGlyValLysThrLysGlnIleArgLysArgValValArgValPheGlnSerGly 320
DB 1077 GTATATGGAGTTAAGACCAACAGATCCGTAAGAGTTGTGAGGCTGTTTCAAGTGGG 1136

QY 321 GlnGlyMetGlyIleLysAlaSerGlu 329
DB 1137 CAGGGAATGGCATCAAGGCATCTGAG 1163

RESULT 8

US-10-292-798-797
; Sequence 797, Application US/10292798
; Publication No. US2003023583A1
; GENERAL INFORMATION:
; APPLICANT: SUWA, MAKIKO
; APPLICANT: ASAI, KIYOSHI
; APPLICANT: AKIYAMA, YUTAKA
; APPLICANT: ABURATANI, HIROYUKI
; TITLE OF INVENTION: GUANOSINE TRIPHOSPHATE-BINDING PROTEIN COUPLED RECEPTORS
; FILE REFERENCE: 084335/166
; CURRENT APPLICATION NUMBER: US/10/292,798
; CURRENT FILING DATE: 2002-11-13
; PRIOR FILING DATE: 2001-11-13
; PRIOR FILING DATE: 2001-12-18
; PRIOR APPLICATION NUMBER: JP 2001-246789
; NUMBER OF SEQ ID NOS: 2070
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 797
; LENGTH: 1390
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; LOCATION: source
; FEATURE:
; LOCATION: (1) .. (1390)
; NAME/KEY: CDS
; LOCATION: (201) .. (1190)
US-10-292-798-797

Alignment Scores:
Pred. No.: 2.05e-192 Length: 1390
Score: 1691.00 Matches: 325
Percent Similarity: 99.09% Conservatives: 1
Best Local Similarity: 98.78% Mismatches: 3
Query Match: 98.43% Indels: 0
DB: 16 Gaps: 0

US-10-081-775-2 (1-329) x US-10-292-798-797 (1-1390)

QY 1 MetSerSerThrLeuGlyHisAsnMetGluSerProHisHisThrAspValAspProSer 20
DB 201 ATGTCAGACACTCTTGGCCACACATGGAATCTCTTAATCACACTGATGTGACCTTCT 260
QY 21 ValPhePheLeuLeuGlyIleProGlyLeuGluGlnPheHisLeuTrpLeuSerLeuPro 40
DB 261 GTCTTCTCTCTGGGCATCCAGGCTCGGAACAATTTCAATTTGGGCTCTCACTCCCT 320
QY 41 ValCysGlyLeuGlyThrAlaThrIleValGlyAsnIleThrIleLeuValValAla 60
DB 321 GTGTGTGGCTTAGGCACAGCCCAATTTGGGCAATATAACTATCTTGTGTTGTTGCC 380
QY 61 ThrGluProValLeuHisLysProValTyrLeuPheLeuCysMetLeuSerThrIleAsp 80
DB 381 ACTGAACAGCTTTGCACAAAGCTGTGTACCTTTTCTGTGATGCTCTCAACCATCGAC 440
QY 81 LeuAlaAlaSerValSerThrValProLysLeuLeuAlaIlePheTrpCysGlyAlaGly 100
DB 441 TTGGCTGCTCTGTCTCCACAGTTCACAGTTCACAGTCTGCTATCTTCTGGTGTGGAGCCGGA 500
QY 101 HisIleSerAlaSerAlaCysLeuAlaHisMetPhePheIleHisAlaPheCysMetMet 120
DB 501 CATATATCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 560
QY 121 GluSerThrValLeuLeuAlaMetAlaPheAspArgTyrValAlaIleCysHisProLeu 140
DB 561 GAGTCCACTGTGCTACTGGCCATGGGCTTTGTATCGCTAGCTGCGCATCTGCCACCATCT 620

```
Qy 141 ArgTyrAlaThrIleLeuThrAspThrIleleAlaHisIleGlyValAlaAlaValVal 160
Db 621 CGCTATGCGCAAACTCTCACTGACACCATCATTTGCCACATAGGGGTGGCAGCTGTAGTG 680
Qy 161 ArgGlySerLeuLeuMetLeuProCysPhePheIleGlyArgLeuAsnPheCysGln 180
Db 681 CGAGGCTCCCTCGCTCATGCTCCCATGTCCCTTCTTATTTGGGCGGTTTGAACCTTTGCCAA 740
Qy 181 SerHisValIleLeuHisThrTyrCysGluHisMetAlaValValLeuLeuAlaCysGly 200
Db 741 AGCCATGTGATCCACACAGTACTGTGACACATGGCTGTGGTGAAGCTGGCCTGTGGA 800
Qy 201 AspThrArgProAsnArgValTyrGlyLeuThrAlaAlaLeuLeuValIleGlyValAsp 220
Db 801 GACACAGGCTTAACCGTGTGTATGGGCTGACAGCTGCACCTGTGGTCAATTGGGGTTGAC 860
Qy 221 LeuPheCysIleGlyLeuSerTyrAlaLeuIleAlaGlnAlaValLeuArgLeuSerSer 240
Db 861 TTGTTTTGCATTTGGTCTCTCTCTATGCCCTTAAGTGCACAGCTGTCTTGGCCTCTCATCC 920
Qy 241 HisGluAlaArgSerLysAlaLeuGlyThrCysGlySerHisValCysValIleLeuIle 260
Db 921 CATGAAGCTCGTCCAGGCCCTAGGACCTGTGGTTCCTCATGTCTGTCATCTCATC 980
Qy 261 SerTyrThrProAlaLeuPheSerPhePheThrHisArgPheGlyHisHisValProVal 280
Db 981 TCTTATACACACCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1040
Qy 281 HisIleHisIleLeuLeuAlaAsnValTyrLeuLeuLeuProProAlaLeuAsnProVal 300
Db 1041 CATATTCACATTCCTTTGGCCAAATGTTTATCTGCTTTTGGCCACCTGCTCTTAACTCTGTG 1100
Qy 301 ValTyrGlyValIleGlyThrIleArgLysArgValValArgValPheGlnSerGly 320
Db 1101 GTATATGGAGTTAAGACCAACAGATCCGTAAGAGAGTTGTTCAGGGTGTTCCTCAAGTGGG 1160
Qy 321 GlnGlyMetGlyIleLysAlaSerGlu 329
Db 1161 CAGGGAATGGCATCAAGGCATCTGAG 1187

RESULT 9
US-10-297-021-42
; Sequence 42, Application US/10297021
; Publication NO. US20040023294A1
; GENERAL INFORMATION:
; APPLICANT: INCYTE GENOMICS, INC.
; APPLICANT: PATTERSON, Chandra
; APPLICANT: TRIBOULEY, Catherine M.
; APPLICANT: YAO, Monique G.
; APPLICANT: GRIFFIN, Jennifer A.
; APPLICANT: THORNTON, Michael
; APPLICANT: LU, Yan
; APPLICANT: KALILICK, Deborah A.
; APPLICANT: GANDHI, Ameena R.
; APPLICANT: AU-YOUNG, Janice
; TITLE OF INVENTION: G-PROTEIN COUPLED RECEPTORS
; FILE REFERENCE: SF-0781 PCT
; CURRENT APPLICATION NUMBER: US/10/297,021
; CURRENT FILING DATE: 2001-05-22
; PRIOR APPLICATION NUMBER: 60/206,222; 60/207,476; 60/208,834; 60/208,861; 60/209,868
; PRIOR FILING DATE: 2000-05-22; 2000-05-25; 2000-06-02; 2000-06-02; 2000-06-07
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: PERL Program
; SEQ ID NO 42
; LENGTH: 966
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. US20040023294A1 7476077CB1
US-10-297-021-42
```

Alignment Scores:

```
Pred. No.: 1,28e-187 Length: 966
Score: 1649.00 Matches: 317
Percent Similarity: 99.07% Conservative: 1
Best Local Similarity: 98.75% Mismatches: 3
Query Match: 95.98% Indels: 0
DB: 17 Gaps: 0

US-10-081-775-2 (1-329) x US-10-297-021-42 (1-966)

Qy 9 MetGluSerProHisHisThrAspValAspProSerValPhePheLeuLeuGlyIlePro 28
Db 1 ATGGAATCTCTAATCAACATGATTTGACCTTCTGTCTTCTCTCTCTCTCTCTCTCTCT 60
Qy 29 GlyLeuGluGlnPheHisLeuTrpLeuSerLeuProValCysGlyLeuGlyThrAlaThr 48
Db 61 GGTCTGGAAACAATTTCAATTTGTGGCTCTCACTCCCTGTGTGGCTTAGGCACAGCCACA 120
Qy 49 IleValGlyAsnIleThrIleLeuValValAlaThrGluProValLeuHisLysPro 68
Db 121 ATGTGGGCAATATAACTATTCTGTGTGTGGCTGAAACCACTGTAACCAAGCTCTTGCAAGGCT 180
Qy 69 ValTyrLeuPheLeuCysMetLeuSerThrIleAspLeuAlaAlaSerValSerThrVal 88
Db 181 GTGTACCTTTTCTGTGCATGCTCTCAACCACTGCACTTGGCTGCTCTGTCTCCACAGTT 240
Qy 89 ProLysLeuLeuAlaIlePheTrpCysGlyAlaGlyHisIleSerAlaSerAlaCysLeu 108
Db 241 CCCAAGCTACTGGCTATCTTCTGGTGTGGAGCCGACATATATCTGCTCTGCTCTGCTGCTG 300
Qy 109 AlaHisMetPhePheIleHisAlaPheCysMetMetGluSerThrValLeuLeuAlaMet 128
Db 301 GCACAGATGTTCTTCAATTCATGCTTGTGATGATGGAGTCCACTGTGTCTACTGGCCATG 360
Qy 129 AlaPheAspArgTyrValAlaIleCysHisProLeuArgTyrAlaThrIleLeuThrAsp 148
Db 361 GCCTTTGATCGCTAGCTGGCCATCTGCCACCCACATCCGCTATGCCACAAATCTCTCACTGAC 420
Qy 149 ThrIleIleAlaHisIleGlyValAlaAlaValArgGlySerLeuLeuMetLeuPro 168
Db 421 ACCATCATGCCCATAGGGGTGGAGCTGTAGTGGAGGGTCTCTGTCTCATGCTCCCA 480
Qy 169 CysProPhePheIleGlyArgLeuAsnPheCysGlnSerHisValIleLeuHisThrTyr 188
Db 481 TGTCTCTCTCTTATTGGGCGTTTGAACCTTCTGCCAAAGCAATGTGTATCTCTACACAGTAC 540
Qy 189 CysGluHisMetAlaValIleValIleCysGlyAspThrArgProAsnArgValTyr 208
Db 541 TGTGAGCACATGGCTGTGGTGAAGCTGTGGAGTGGAGACACAGGCGCTTAACCGTGTGTAT 600
Qy 209 GlyLeuThrAlaAlaLeuLeuValIleGlyValAspLeuPheCysIleGlyLeuSerTyr 228
Db 601 GGGCTGACAGCTGCATGTTGGTTCATTTGGGGTTGACTTGTGTTCATTTGCTCTCTCTAT 660
Qy 229 AlaLeuIleAlaGlnAlaValLeuArgLeuSerSerHisGluAlaAlaArgSerLysAlaLeu 248
Db 661 GCCCTAAGTGCAAGCTGTCTTGTCTCTCATCCATGAAGCTCGTCCAAGGCCCTA 720
Qy 249 GlyThrCysGlySerHisValCysValIleLeuLeuSerTyrThrProAlaLeuPheSer 268
Db 721 GGGACCTGTGGTTCCTCATCTGTGTGTCATCTCTTATACACAGCCCTCTCTCTCTC 780
Qy 269 PhePheThrHisArgPheGlyHisHisValProValHisIleHisIleLeuLeuAlaAsn 288
Db 781 TTTTITACACACCGCTTTGGCCATCACGTTCCAGTCCATATTCACATCTCTTTTGGCCAAT 840
Qy 289 ValTyrLeuLeuLeuProProAlaLeuAsnProValValTyrGlyValIleThrLysGln 308
Db 841 GTTTATCTGCTTTTGGCAGCTGTCTTAATCTCTGTGTGTATATGGAGTTAAGACCAACAG 900
Qy 309 IleArgLysArgValValArgValPheGlnSerGlyGlnGlyMetGlyIleLysAlaSer 328
Db 901 ATCCGTAAGAGAGTTGTTCAGGGTGTTCAGAGTGGGCAAGGCAATGGGCATCAAGGCATCT 960
```


Alignment Scores:		
Pred. No.:	1.03e-103	Length: 1242
Score:	951.50	Matches: 180

Percent Similarity:	73.12%	Conservative:	54
Best Local Similarity:	56.25%	Mismatches:	83
Query Match:	55.38%	Indels:	4
DB:	15	Gaps:	2
US-10-081-775-2 (1-329) x US-10-017-161-959 (1-1242)			
Qy	1	MetSerSerThrLeuGlyHisAsnMetGluSerProHisHisThrAspValAspProSer	20
Db	78	CTAACCACTTACTTGGCCATGAACACTCATAAACCATACCATCAG-----AACCCCAACC	131
Qy	21	ValPhePheLeuGlyIleProGlyLeuGluGlnPheHisLeuThrLeuSerLeuPro	40
Db	132	TCCTTTCTGCTCATGGGAATTCAGGCCCGGAGGCAATCCCACTTTTGGATTGCTTTCC	191
Qy	41	ValCysGlyLeuGlyThrAlaThrIleValGlyAsnIleThrIleLeuValValAla	60
Db	192	TTCTGCTCCATGATGCCCTGGCAGTGTCTGGAAACATGGTGTCTGCTAGTGTATCAT	251
Qy	61	ThrGluProValLeuHisGlyProValTyrIlePheLeuCysMetLeuSerThrIleAsp	80
Db	252	TCAGAGCTGTATTGACACAGCCCATGTACCTGTCTGATGCTATCCACCATTCAC	311
Qy	81	LeuAlaAlaSerValSerThrValProLysLeuLeuAlaIlePheTyrCysGlyAlaGly	100
Db	312	CTGGTCTCTGCACCTCCACTGTGCCAAGCTCTTGCACCTTTTGGCAAGGATGCT	371
Qy	101	HisIleSerAlaSerAlaCysLeuAlaHisMetPhePheIleHisAlaPheCysMetMet	120
Db	372	GAGATCAACTTTGGGGCTGTGTGCCAGATGTTCTTATCCATGGCTTCTCAGCTGTA	431
Qy	121	GluSerThrValLeuLeuAlaMetAlaPheAspArgTyrValAlaIleCysHisProLeu	140
Db	432	GAATCTGTATACTCTAGCATGGCTTTGACCGCTACTTAGCCATTTGCTGGCCTCTG	491
Qy	141	ArgTyrAlaThrIleLeuThrAspThrIleAlaHisIleGlyValAlaAlaValVal	160
Db	492	CACATATGGTTCATCTCTCCAGAGCTGTAGGCAAGCTGGGGCTGCAGC-GTGCTT	550
Qy	161	ArgGlySerLeuLeuMetLeuProCysProPhePheIleGlyArgLeuAsnPheCysGln	180
Db	551	CGTGTGTGGGACTCATGCCCACTCACTCTTACTTGGCAAGACTGAGTACTGTC---	607
Qy	181	SerHisValIleLeuHisThrTyrCysGluHisMetAlaValValIleLeuAlaCysGly	200
Db	608	AGTCAGTGTGGCCCACTCTACTGTGAACACATGGCTGTGAAGCTGGCTGTGGA	667
Qy	201	AspThrArgProAsnArgValTyrGlyLeuThrAlaAlaLeuLeuValIleGlyValAsp	220
Db	668	GGAAACACAGCCMAACAATCTATGGCATCTGTGCACACTGGTGTGGGCACTGAC	727
Qy	221	LeuPheCysIleGlyLeuSerTyrAlaLeuIleAlaGlnAlaValLeuArgLeuSerSer	240
Db	728	TCCATCTGTATTGCTCTCTCTATGCACTCATCTCCGAGCTGTGTAGGTCTTTCCTCC	787
Qy	241	HisGluAlaArgSerIleAlaLeuGlyThrCysGlySerHisValCysValIleLeuIle	260
Db	788	AAGGAGGCAAGGGCTAAGACCTTTGGCACTTGTGGCTCCCACTGGGTGTACTTCTC	847
Qy	261	SerTyrThrProAlaLeuPheSerPhePheThrHisArgPheGlyHisValProVal	280
Db	848	TTCTACACACAGGACTCTCTCTCTCTACACACAGCGGTGGCCAGCAGCTGCCCGG	907
Qy	281	HisIleHisIleLeuLeuAlaAsnValTyrLeuLeuLeuProProAlaLeuAsnProVal	300
Db	908	CACATCCATCCTTCTAGCTACCTCTACCTGGTTGTGCCACCCCACTGCTCAACCCCA	967
Qy	301	ValTyrGlyValIleThrLysGlnIleArgLysArgValValArgValPheGlnSerGly	320
Db	968	ATCTATGGCATGAAGAACCAACAGATCTGGGATGGGGCCCTCCGGCTTCTGAAGTGGGC	1027
RESULT 12			
US-10-343-650A-315			

; Sequence 315, Application US/10343650A			
; Publication No. US20040067499A1			
; GENERAL INFORMATION:			
; APPLICANT: HAGA, TATSUYA			
; TITLE OF INVENTION: NOVEL G-PROTEIN COUPLED RECEPTOR			
; FILE REFERENCE: 31671-186347			
; CURRENT APPLICATION NUMBER: US/10/343,650A			
; CURRENT FILING DATE: 2003-07-21			
; PRIOR APPLICATION NUMBER: JP 2000/237818			
; PRIOR FILING DATE: 2000-08-04			
; PRIOR APPLICATION NUMBER: JP 2001/34434			
; PRIOR FILING DATE: 2001-02-13			
; NUMBER OF SEQ ID NOS: 694			
; SOFTWARE: PatentIn Ver. 2.1			
; SEQ ID NO 315			
; LENGTH: 990			
; TYPE: DNA			
; ORGANISM: Homo sapiens			
; FEATURE:			
; NAME/KEY: CDS			
; LOCATION: (1)..(990)			
US-10-343-650A-315			
Alignment Scores:			
Pred. No.:	1.81e-101	Length:	990
Score:	931.50	Matches:	179
Percent Similarity:	73.91%	Conservative:	42
Best Local Similarity:	59.87%	Mismatches:	77
Query Match:	54.22%	Indels:	1
DB:	13	Gaps:	1
US-10-081-775-2 (1-329) x US-10-343-650A-315 (1-990)			
Qy	19	ProSerValPhePheLeuGlyIleProGlyLeuGluGlnPheHisLeuThrLeuSer	38
Db	82	CTTCTCT---TTCTCTGTAGGGATTCCAGTTTAGAGGAAGCAGCAGCTGGATTGCA	138
Qy	39	LeuProValCysGlyLeuGlyThrAlaThrIleValGlyAsnIleThrIleLeuValVal	58
Db	139	CTGCCCTGGGCATCTTTTACCTCTCTTTAGTGGCAATGTTTACCATTCTCTTCATC	198
Qy	59	ValAlaThrGluProValLeuHisLysProValTyrLeuPheLeuCysMetLeuSerThr	78
Db	199	ATCTGGATGGACCCATCTCTTGACCAATCTATGTACCTCTTCTCTCCATGTAGCTGCC	258
Qy	79	IleAspLeuAlaAlaSerValSerThrValProLysLeuLeuAlaIlePheTyrCysGly	98
Db	259	ATCGACCTGGTTCCTGGCTCTCCACTGCACCCAAAGCCCTTGCAGTGTCTCTGTTTCA	318
Qy	99	AlaGlyHisIleSerAlaSerAlaCysLeuAlaHisMetPheIleHisAlaPheCys	118
Db	319	GCCACAGAGATTGGGTACATCTCTGCTGATCCAGATGTTCTTCATCCATGCACTTCTCC	378
Qy	119	MetMetGluSerThrValLeuLeuAlaMetAlaPheAspArgTyrValAlaIleCysHis	138
Db	379	TCCATGGAGTCAGGGGTACTTGTGGCCATGGCTCTGGATCGCTATGTAGCCATTTGTCAC	438
Qy	139	ProLeuArgTyrAlaThrIleLeuThrAspThrIleIleAlaHisIleGlyValAlaAla	158
Db	439	CCCTTGGACCATTCACCAATCTCGATCCAGGGGTCATAGGGCGCATCGGAATGGTGTG	498
Qy	159	ValValArgGlySerLeuLeuMetLeuProCysProPhePheIleGlyArgLeuAsnPhe	178
Db	499	CTGTGTGGGGATTACTACTCTTATCCCTTCCCATTTTGTGGGAACACTTATATCTTC	558
Qy	179	CysGlnSerHisValIleLeuHisThrTyrCysGluHisMetAlaValValIleCysHis	198
Db	559	TGCCAAGCCACCATCATAGGCCATGCCTATTGTGAACATATGCTCTTGTGAACATGCC	618
Qy	199	CysGlyAspThrArgProAsnArgValTyrGlyLeuThrAlaAlaLeuValIleGly	218
Db	619	TGCTCAGAAACACACAGTCAATCGAGCTTATGGGTGACTATGGCTTGTGTGATTGGG	678

```

RESULT 14
US-10-467-252-93
; Sequence 93, Application US/10467252
; Publication No. US20040115676A1
; GENERAL INFORMATION:
; APPLICANT: BAUGHN, Mariah R.; TRIBOULEY, Catherine M.;
; APPLICANT: NGUYEN, Daniel B.; THORNTON, Michael;
; APPLICANT: YAO, Monique G.; KALLICK, Deborah A.;
; APPLICANT: GANDEHI, Ameena R.; CHAWLA, Narinder K.;
; APPLICANT: ARVIZU, Chandra S.; ELLIOTT, Vicki S.;
; APPLICANT: HAFALIA, April J.A.; RAMKUMAR, Jayalaxmi;
; APPLICANT: JIN, Pei; TANG, Y.Tom;
; APPLICANT: YUE, Henry; REDDY, Roopa
; APPLICANT: BURFORD, Neil; LU, Dyung Aina M.;
; APPLICANT: GRAUL, Richard C.; KAHN, Farrah A.;
; APPLICANT: WALSH, Roderick T.; ISON, Craig H.;
; APPLICANT: RICHARDSON, Thomas W.; GRIFFIN, Jennifer A.
; APPLICANT: WARREN, Bridget A.; YANG, Junning;
; APPLICANT: LEE, Ernestine A.; HARLAND, Lee
; TITLE OF INVENTION: G-PROTEIN COUPLED RECEPTORS
; FILE REFERENCE: PI-0357 USN
; CURRENT APPLICATION NUMBER: US/10/467,252
; CURRENT FILING DATE: 2003-08-06
; PRIOR APPLICATION NUMBER: PCT/US02/03635
; PRIOR FILING DATE: 2002-02-06
; PRIOR APPLICATION NUMBER: US 60/267,322;
; PRIOR FILING DATE: 2001-02-07
; PRIOR APPLICATION NUMBER: US 60/271,215
; PRIOR FILING DATE: 2001-02-23
; PRIOR APPLICATION NUMBER: US 60/274,551
; PRIOR FILING DATE: 2001-03-08
; PRIOR APPLICATION NUMBER: US 60/278,507
; PRIOR FILING DATE: 2001-03-23

```

```
; PRIOR APPLICATION NUMBER: US 60/280,597
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: US 60/281,107
; PRIOR FILING DATE: 2001-04-02
; PRIOR APPLICATION NUMBER: US 60/282,121
; PRIOR FILING DATE: 2001-04-06
; NUMBER OF SEQ ID NOS: 96
; SOFTWARE: PERL Program
; SEQ ID NO 93
; LENGTH: 1114
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No: 7475226CB1
US-10-467-252-93

Alignment Scores:
Pred. No.: 2,18e-101 Length: 1114
Score: 931.50 Matches: 179
Percent Similarity: 73.91% Conservative: 42
Best Local Similarity: 59.87% Mismatches: 77
Query Match: 54.22% Indels: 1
DB: 17 Gaps: 1

US-10-081-775-2 (1-329) x US-10-467-252-93 (1-1114)
QY 19 ProSerValPhePheLeuLeuGlyLeuProGlyLeuGluGlnPheHisLeuTrpLeuSer 38
DB 106 CTTCTCTTCTGTTAGGATTCAGGATTTAGAGGAAGCCAGCAGCAGTGGATTGCA 162
QY 39 LeuProValCysGlyLeuGlyThrAlaThrIleValGlyAsnIleThrIleLeuVal 58
DB 163 CTGCCCTGGGATCTTACCTCTGCTTTAGTGGCAATGTTACCAATCTCTTCATC 222
QY 59 ValAlaThrGluProValLeuHisLysProValTyrLeuPheLeuCysMetLeuSerThr 78
DB 223 ATCTGGATGGACCATCTCTGACCAATCTATGTAACCTCTCTGTCATGTCAGTGC 282
QY 79 IleAspLeuAlaAlaSerValSerThrValProLysLeuLeuAlaIlePheTrpCysGly 98
DB 283 ATCGACCTGGTCTGGCTCTCCACTGCACCCAAAGCCCTTGCAGTGCCTCTGGTTCAT 342
QY 99 AlaGlyHisIleSerAlaSerAlaCysLeuAlaHisMetPhePheIleHisAlaPheCys 118
DB 343 GCCCAGCAGATTGGGTACATCGCTGCTGATCCAGATGTTCTTCATCCATGCAATCTCC 402
QY 119 MetMetGluSerThrValLeuLeuAlaMetAlaPheAspArgTyrValAlaIleCysHis 138
DB 403 TCCATGGAGTCAGGGTACTTGGGCCATGGCTCGATCGCTATGATGCCATTGTCAC 462
QY 139 ProLeuArgTyrAlaThrIleLeuThrAspThrIleIleAlaHisIleGlyValAlaAla 158
DB 463 CCCTTGACCAATCCACAATCTGCATCCAGGGGTCTATAGGGCGCATAGGATGGTGGTG 522
QY 159 ValValArgGlySerLeuLeuMetLeuProCysProPhePheIleGlyArgLeuAsnPhe 178
DB 523 CTGGTGGAGGGATTAATCTATCTCTTATCCCTTCCCTTCTGTTGGGAACACTTATCTTC 582
QY 179 CysGlnSerHisValIleLeuHisThrTyrCysGluHisMetAlaValValLysLeuAla 198
DB 583 TGCCAGGCCACCATCATAGCCATGCTATTGGAACATATGCTGTTGGAACACTTATCTTC 642
QY 199 CysGlyAspThrArgProLeuArgValTyrGlyLeuThrAlaAlaIleLeuValIleGly 218
DB 643 TGCTCAGAAACCAACAGTCAATCGAGCTTATGGGCTGACTATGGCTTGTGTTGTTGGG 702
QY 219 ValAspLeuPheCysIleGlyLeuSerTyrAlaLeuIleAlaGlnAlaValLeuArgLeu 238
DB 703 CTGGATGTTCTGGCCATTTGTTCTTATGCCCCACATCTCCAGGAGTCTGTAAGGTA 762
QY 239 SerSerHisGluAlaArgSerLysAlaLeuGlyThrCysGlySerHisValCysValIle 258
```

763 CCAGGAGTGGAGCCCGACCTTAGCGCTTAGCACATGTGGCTCTCATATTTGTGTGTCATC 822

259 LeuIleSerTyrThrProAlaLeuPheSerPhePheThrHisArgPheGlyHisHisVal 278

823 CTGGTCTTCTATGTCCTCGGAAATTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 882

279 ProValHisIleHisIleLeuLeuAlaAsnValTyrLeuLeuLeuProProAlaLeuAsn 298

883 CCCCATCATGTCATGTCCT 942

299 ProValValTyrGlyValLysThrLysGlnIleArgLysArgValValArgValPhe 317

943 CTTCTTCTATGAGTGAAGACTCAGCAGATCCGCCAGCGAGTGTCTCAGAGTGTGT 999

RESULT 15

US-10-017-161-945

; Sequence 945, Application US/10017161

; Publication No. US20030143668A1

; GENERAL INFORMATION:

; APPLICANT: SUMA, MAKIKO

; APPLICANT: ASAI, KIYOSHI

; APPLICANT: AKIYAMA, YUTAKA

; APPLICANT: ABURATANI, HIROYUKI

; TITLE OF INVENTION: NOVEL G PROTEIN-COUPLED RECEPTORS

; FILE REFERENCE: 084335/0152

; CURRENT APPLICATION NUMBER: US/10/017,161

; CURRENT FILING DATE: 2002-12-18

; PRIOR APPLICATION NUMBER: JP 2001/246789

; PRIOR FILING DATE: 2001-06-18

; NUMBER OF SEQ ID NOS: 2430

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 945

; LENGTH: 1390

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: source

; LOCATION: (1)..(1390)

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (201)..(1190)

; FEATURE:

; NAME/KEY: modified base

; LOCATION: (1)..(26)

; OTHER INFORMATION: a, t, c, g, unknown or other

US-10-017-161-945

Alignment Scores:

Pred. No.: 3,07e-101 Length: 1390

Score: 931.50 Matches: 179

Percent Similarity: 73.91% Conservative: 42

Best Local Similarity: 59.87% Mismatches: 77

Query Match: 54.22% Indels: 1

DB: 15 Gaps: 1

US-10-081-775-2 (1-329) x US-10-017-161-945 (1-1390)

QY 19 ProSerValPhePheLeuLeuGlyLeuProGlyLeuGluGlnPheHisLeuTrpLeuSer 38

282 CTTCTCTTCTGTTAGGATTCAGGATTTAGAGGAAGCCAGCAGCAGTGGATTGCA 338

39 LeuProValCysGlyLeuGlyThrAlaThrIleValGlyAsnIleThrIleLeuVal 58

339 CTGCCCTGGGATCTTACCTCTGCTTTAGTGGCAATGTTACCAATCTCTTCATC 398

59 ValAlaThrGluProValLeuHisLysProValTyrLeuPheLeuCysMetLeuSerThr 78

399 ATCTGGATGGACCAATCTATGTAACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 458

79 IleAspLeuAlaAlaSerValSerThrValProLysLeuLeuAlaIlePheTrpCysGly 98

459 ATCGACCTGGTCTGGCT 518

QY 99 AlaGlyHisIleSerAlaSerAlaCysLeuAlaHisMetPhePheIleHisAlaPheCys 118
Db |||||
519 GCCACGAGATTGGGTACATCTCTGCTGATCCAGATGTTCTTCATCCATGCAITCTCC 578
QY 119 MetMetGluSerThrValLeuLeuAlaMetAlaPheAspArgTyrValAlaIleCysHis 138
Db |||||
579 TCCATGGAGTCAGGGGTACTTGTGGCCATGGCTCTGGATCGTATGTAGCCATTGTGCAC 638
QY 139 ProLeuArgTyrAlaThrIleLeuThrAspThrIleIleAlaHisIleGlyValAlaAla 158
Db |||||
639 CCTTGCACCATTCACAACTCTGCATCCAGGGGTCATAGGCGCATCGGAATGGTGGTG 698
QY 159 ValValArgGlySerLeuLeuMetLeuProCysProPhePheIleGlyArgLeuAsnPhe 178
Db |||||
699 CTGGTGAGGGATTACTACTCTATCCCTTCCCATTTTGTGGGAACATTATCTTC 758
QY 179 CysGlnSerHisValIleLeuHisThrTyrCysGluHisMetAlaValValIysLeuAla 198
Db |||||
759 TGCCAAGCCACCATCATAGGCCATGCTATTGTGAACATATGGCTGTGTGAACCTTGCC 818
QY 199 CysGlyAspThrArgProAsnArgValTyrGlyLeuThrAlaAlaLeuLeuValIleGly 218
Db |||||
819 TGCTCAGAAACACAGTCAATCGAGCTTATGGGCTGACTATGGCCTTGTGTGATTGGG 878
QY 219 ValAspLeuPheCysIleGlyLeuSerTyrAlaLeuIleAlaGlnAlaValLeuArgLeu 238
Db |||||
879 CTGGATGTTCTGGCCATTGGTGTTCCTATGCCACATCCTCCAGGCAGTCTGAAGTA 938
QY 239 SerSerHisGluAlaArgSerIysAlaLeuGlyThrCysGlySerHisValCysValIle 258
Db |||||
939 CCAGGGAGTGAGGCCCGACTTAAGGCGTTTAGCACATGTGGCTCTCATATTGTGTCTATC 998
QY 259 LeuIleSerTyrThrProAlaLeuPheSerPhePheThrHisArgPheGlyHisVal 278
Db |||||
999 CTGTCCTCTATGTCCTCGGAATTTCTCTCTCACTACCCGCTTTGGTCAATCATGTA 1058
QY 279 ProValHisIleHisIleLeuLeuAlaAsnValTyrLeuLeuLeuProProAlaLeuAsn 298
Db |||||
1059 CCCCATCATGTCCATGTTCTTCTGGCCACACGGTATCTCTCATGCCACCTGGCGTCAAT 1118
QY 299 ProValValTyrGlyValIysGlnIleArgLysArgValValArgValPhe 317
Db |||||
1119 CCTCTGTCTATGGAGTGAGACTCAGCAGATCCGCCAGCGAGTGTCTCAGAGTGTGT 1175

Search completed: August 27, 2004, 21:07:04
Job time : 556 secs

This page blank (b)(7)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: August 27, 2004, 18:30:23 ; Search time 90 Seconds

(without alignments)
2028.654 Million cell updates/sec

Title: US-10-081-775-2

Perfect score: 1718

Sequence: 1 MSSTLGHMESPHHTVDPS.....RKRVRVYFQSGQGMKASE 329

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO.spool/US10081775/runat_20082004_170214_11153/app_query.fasta_1.519
-DB=Issued Patents NA -QFMT=fastap -SUFFIX=rni -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bites -START=1 -END=1 -MATRIX=blomsum62 -TRANS=human40.cdd
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10081775@cgn1_1_69@runat_20082004_170214_11153 -NCPU=6 -ICPU=3
-NO_WMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents NA.*

1: /cgn2_6/ptodata/2/ina/5A.COMB.seq.*
2: /cgn2_6/ptodata/2/ina/5B.COMB.seq.*
3: /cgn2_6/ptodata/2/ina/6A.COMB.seq.*
4: /cgn2_6/ptodata/2/ina/6B.COMB.seq.*
5: /cgn2_6/ptodata/2/ina/PCTUS.COMB.seq.*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	867.5	50.5	1539	US-09-668-680-13	Sequence 13, Appl
2	861	50.1	963	US-09-439-313-526	Sequence 526, App
3	861	50.1	963	US-09-636-215-526	Sequence 526, App
4	861	50.1	963	US-09-685-166A-526	Sequence 526, App
5	856	49.8	1474	US-08-465-980-1	Sequence 1, Appli
6	856	49.8	1474	US-09-053-303-1	Sequence 1, Appli
7	856	49.8	1474	US-09-339-115-1	Sequence 1, Appli
8	856	49.8	1474	PCT-US95-07093-1	Sequence 1, Appli
9	424	24.7	966	US-08-748-506-5	Sequence 5, Appli
10	423.5	24.7	966	US-08-748-506-7	Sequence 7, Appli
11	416.5	24.2	966	US-08-748-506-6	Sequence 6, Appli
12	410.5	23.9	966	US-08-748-506-8	Sequence 8, Appli

13	410	23.9	1282	4	US-09-016-434-1413	Sequence 1413, Ap
14	410	23.9	1351	4	US-09-546-986A-5	Sequence 5, Appli
15	410	23.9	1351	4	US-09-524-730-5	Sequence 5, Appli
16	404.5	23.5	1438	4	US-09-016-434-1313	Sequence 1313, Ap
17	401	23.3	1035	4	US-09-546-986A-1	Sequence 1, Appli
18	401	23.3	1035	4	US-09-524-730-1	Sequence 1, Appli
19	398.5	23.2	1854	4	US-09-016-434-1312	Sequence 1312, Ap
20	396.5	23.1	1713	2	US-08-467-948A-1	Sequence 1, Appli
21	396.5	23.1	1713	3	US-08-467-947A-1	Sequence 1, Appli
22	390	22.7	1062	4	US-09-668-680-10	Sequence 10, Appl
23	387	22.5	945	4	US-09-016-434-1114	Sequence 1114, Ap
24	385	22.4	1990	4	US-09-016-434-1056	Sequence 1056, Ap
25	381.5	22.2	900	3	US-09-085-371-5	Sequence 5, Appli
26	381	22.2	1065	4	US-09-546-986A-7	Sequence 7, Appli
27	381	22.2	1065	4	US-09-524-730-7	Sequence 7, Appli
28	376.5	21.9	3459	4	US-09-016-434-1363	Sequence 1363, Ap
29	374	21.8	1290	2	US-08-827-291A-1	Sequence 1, Appli
30	373.5	21.7	952	4	US-09-016-434-1115	Sequence 1115, Ap
31	352	20.5	1080	4	US-09-668-680-9	Sequence 9, Appli
32	348.5	20.3	1297	4	US-09-668-680-11	Sequence 11, Appl
33	325	18.9	984	3	US-08-748-506-9	Sequence 9, Appli
34	320	18.6	951	4	US-09-465-901-47	Sequence 47, Appl
35	311.5	18.1	1411	4	US-09-546-986A-3	Sequence 3, Appli
36	311.5	18.1	1411	4	US-09-524-730-3	Sequence 3, Appli
37	310.5	18.1	666	4	US-09-465-901-33	Sequence 33, Appl
38	305	17.8	678	4	US-09-465-901-45	Sequence 45, Appl
39	299.5	17.4	675	4	US-09-465-901-43	Sequence 43, Appl
40	297	17.3	675	4	US-09-465-901-17	Sequence 17, Appl
41	293	17.1	669	4	US-09-465-901-37	Sequence 37, Appl
42	289	16.8	669	4	US-09-465-901-11	Sequence 11, Appl
43	287	16.7	669	4	US-09-465-901-35	Sequence 35, Appl
44	285	16.6	669	4	US-09-465-901-39	Sequence 39, Appl
45	283	16.5	669	4	US-09-465-901-15	Sequence 15, Appl

ALIGNMENTS

RESULT 1
US-09-668-680-13
; Sequence 13, Application US/09668680
; Patent No. 6436703
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Zhou, Ping
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhang, Jie
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Xue, Aidong J.
; APPLICANT: Xu, Chongjun
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. 6436703el Nucleic Acids and
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 790CIP2A
; CURRENT APPLICATION NUMBER: US/09/668,680
; CURRENT FILING DATE: 2000-09-22
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: pt_FL_genes Version 2.0
; SEQ ID NO 13
; LENGTH: 1539
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (130)..(1539)
US-09-668-680-13
Alignment Scores: 1.04e-92 Length: 1539
Pred. No.: 1.04e-92 Length: 1539


```
Score: 867.50 Matches: 156
Percent Similarity: 69.70% Conservative: 51
Best Local Similarity: 52.53% Mismatches: 89
Query Match: 50.49% Indels: 1
DB: 4 Gaps: 1

US-10-081-775-2 (1-329) x US-09-668-680-13 (1-1539)

Qy 26 GlytLeuProGlyLeuGluGlnPheHisLeuTrpLeuSerLeuProValCysGlyLeuGly 45
Db 637 GGGATCCAGGGCTGGAGGCTGCCACCTTCGGATTGCCATCCCTTTCTGGCCATGTAT 696
Qy 46 ThrAlaThrIleValGlyAsnIleThrIleLeuValValAlaThrGluProValLeu 65
Db 697 CTTGTAGACCTGGTGGAAATGTCGCCCTCATCTGCTGCTATGTCATGCGACAATGCTCTT 756
Qy 66 HisLysProValTyrLeuPheLeuCysMetLeuSerThrIleAspLeuAlaAlaSerVal 85
Db 757 CATGCACCTATGTACTCTTCCTGCTGCTTCTCTCACTCACAGACCTGGCTCTCAGTTCT 816
Qy 86 SerThrValProLysLeuLeuAlaIlePheTrpCysGlyAlaGlyHisIleSerAlaSer 105
Db 817 ACCACTGTGCCAAGATGCTGGCCATTTTGTGGCTCCATGCTGTGATTCCTTTGGT 876
Qy 106 AlaCysLeuAlaHisMetPhePheIleHisAlaPheCysMetMetGluSerThrValLeu 125
Db 877 GGATGCTGGCCAGATGTTTGTGTCATTTCTATCTATGCTCTGGAGTCTCTCGATTCTA 936
Qy 126 LeuAlaMetAlaPheAspArgTyrValAlaIleCysHisProLeuArgTyrAlaThrIle 145
Db 937 CTTGCCATGGCTTTGATAGGTATGCTGCTATCTGTAACCCATTAAGGTATACAAACCAT 996
Qy 146 LeuThrAspThrIleIleAlaHisIleGlyValAlaValAlaValArgGlySerLeuLeu 165
Db 997 CTCAACCATGCTGTCATAGGAGGAGATTTGCTGGCTATTCCTGCTAGTGGCTATT 1056
Qy 166 MetLeuProCysProPhePheIleGlyArgLeuAsnPheCysGlnSerHisValIleLeu 185
Db 1057 GTCTCCCTTCATCTTCTGCTGAGGAGATCCCTCTACTGTGCTACCGTGTGATGACA 1116
Qy 186 HisThrTyrCysGluHisMetAlaValValLysLeuAlaCysGlyAspThrArgProAsn 205
Db 1117 CACACATCTGTGAGCATATGGCATCGCCGAGCTGGCTGTGCCAATCACTGTCAAT 1176
Qy 206 ArgValTyrGlyLeuThrAlaAlaLeuValIleGlyValAspLeuPheCysIleGly 225
Db 1177 ATTGCTATGGCTTAACCTGTGGCTCTGCTGCCATGGGACTGGATTCCATTTCTCATTTGCC 1236
Qy 226 LeuSerTyrAlaLeuIleAlaGlnAlaValLeuArgLeuSerHisGluAlaArgSer 245
Db 1237 ATTTCTATGGCTTTATCTCTCATGCTGCTTTTCCATTTCCATCTCATGATGCCAGCAC 1296
Qy 246 LysAlaLeuGlyThrCysGlySerHisValCysValIleLeuIleSerTyrThrProAla 265
Db 1297 AAAGCTCTGAGTACCTGTGGCTCCACATTTGGCATCATCTGGTTTCTACATCCCTGCC 1356
Qy 266 LeuPheSerPhePheThrHisArgPheGlyHisHis---ValProValHisIleHisIle 284
Db 1357 TTCTTCTCTTCTCCTCACCCACCGCTTTGGTCACCAAGGAGTCCCAAGCATGTGCACATC 1416
Qy 285 LeuLeuAlaAsnValTyrLeuLeuProProAlaLeuAsnProValValTyrGlyVal 304
Db 1417 TTTTGGCTATCTATGCTGCTGGTCTCTCTGTACTCATCTCTATTTCTATGGAGCT 1476
Qy 305 LysThrLysGlnIleArgLysArgValValArgValPheGlnSerGlyGln 321
Db 1477 AGAACCAAGGAGATTCGGAGTCGACTTCTAAACTGCTTCACTCTTCACTCTGCTGCTG 1527

RESULT 2
US-09-439-313-526
; Sequence 526, Application US/09439313
; Patent No. 6329505
; GENERAL INFORMATION:
```

```
APPLICANT: Xu, Jiangchun
APPLICANT: Dillon, Davin C.
APPLICANT: Mitcham, Jennifer L.
APPLICANT: Harlocker, Susan Louise
APPLICANT: Jiang Yuqi
APPLICANT: Reed, Steven G.
APPLICANT: Kalos, Michael
APPLICANT: Fanger, Gary
APPLICANT: Retter, Mark
APPLICANT: Solk, John
APPLICANT: Day, Craig
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND
FILE REFERENCE: 210121.427C9
CURRENT APPLICATION NUMBER: US/09/439,313
CURRENT FILING DATE: 1999-11-12
NUMBER OF SEQ ID NOS: 575
SOFTWARE: Fast-Seq for Windows Version 3.0
SEQ ID NO 526
LENGTH: 963
TYPE: DNA
ORGANISM: Homo sapiens
US-09-439-313-526

Alignment Scores:
Pred. No.: 2,95e-92 Length: 963
Score: 861.00 Matches: 161
Percent Similarity: 70.23% Conservative: 49
Best Local Similarity: 53.85% Mismatches: 89
Query Match: 50.12% Indels: 0
DB: 4 Gaps: 0

US-10-081-775-2 (1-329) x US-09-439-313-526 (1-963)

Qy 20 SerValPhePheLeuLeuGlyIleProGlyLeuGluGlnPheHisLeuTrpLeuSerLeu 39
Db 25 GCCACCTTTGGCTTATTGGTATCCAGGATTAGAGAAAGCCCATTTCTGGTTGGCTTC 84
Qy 40 ProValCysGlyLeuGlyThrAlaThrIleValGlyAsnIleThrIleLeuValValVal 59
Db 85 CCCCTCTCTTCATGTATGATGGCAATGTTGGAACTGTCATCGTGGTCTTCATCGTA 144
Qy 60 AlaThrGluProValLeuHisLysProValTyrLeuPheLeuCysMetLeuSerThrIle 79
Db 145 AGGACGGAACGACACCTGCGCTCCGATGACCTCTTCTCTGCACTGCTGACGCCATT 204
Qy 80 AspLeuAlaAlaSerValSerThrValProLysLeuLeuAlaIlePheTyrCysGlyAla 99
Db 205 GACCTGGCTTATCCACATCCACCATGCTAAGATCCTTGCCCTTTCTGCTTTGATTTC 264
Qy 100 GlyHisIleSerAlaSerAlaCysLeuAlaHisMetPhePheIleHisAlaPheCysMet 119
Db 265 CGAGAGATTAGCTTTGAGGCTGCTTACCAGATGTTCTTTATTCATGCCCCCTCTCAGCC 324
Qy 120 MetGluSerThrValLeuLeuAlaMetAlaPheAspArgTyrValAlaIleCysHisPro 139
Db 325 ATTGAATCCACCATCTCTGCGCATGGCTTTGACCGTTATGTGGCATCTGCCACCCA 384
Qy 140 LeuArgTyrAlaThrIleLeuThrAspThrIleIleAlaHisIleGlyValAlaVal 159
Db 385 CTGGCCATGTGCGATGCTCAACATACAGTAACAGCCCAAGATGGCATCGTGGCTGTG 444
Qy 160 ValArgGlySerLeuLeuMetLeuProCysProPhePheIleGlyArgLeuAsnPheCys 179
Db 445 GTCCGGGATTCCTCTTTTTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 504
Qy 180 GlnSerHisValIleLeuHisThrTyrCysGluHisMetAlaValValLysLeuAlaCys 199
Db 505 CACTCCAATGCTCTCGCATCTCTATGTGTCCACAGGATGTAATGAAGTTGGCTTAT 564
Qy 200 GlyAspThrArgProAsnArgValTyrGlyLeuThrAlaAlaLeuLeuValIleGlyVal 219
Db 565 GCAGACACTTTGGCCCAATGTGTATGTATGCTTACTGCACTTCTGCTGCTGCTGCTGCTG 624
```



```
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Hepier, William
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
; FILE REFERENCE: 210121.427C21
; CURRENT APPLICATION NUMBER: US/09/685.166A
; CURRENT FILING DATE: 2000-10-10
; NUMBER OF SEQ ID NOS: 898
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 526
; LENGTH: 963
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-685-166A-526

Alignment Scores:
Pred. No.: 2.95e-92 Length: 963
Score: 861.00 Matches: 161
Percent Similarity: 70.23% Conservative: 49
Best Local Similarity: 53.85% Mismatches: 89
Query Match: 50.12% Indels: 0
DB: 4 Gaps: 0

US-10-081-775-2 (1-329) x US-09-685-166A-526 (1-963)

Qy 20 SerValPhePheLeuLeuGlyProGlyLeuGlnPheHisLeuSerLeu 39
Db 25 GCCACCTTTGGCTTTATTTGGTATCCAGGATCCAGAAAGCCCATTTCTGGGTGGGCTTC 84
Qy 40 ProValCysGlyLeuGlyThrAlaThrIleValGlyAsnIleThrIleLeuValVal 59
Db 85 CCCTCTCTTCATGTATGTAGTGGCATGTTTGGAACTGCATCGTGGTCTTCATCGTA 144
Qy 60 AlaThrGluProValLeuHisLysProValTyrIlePheLeuPheCysMetLeuSerThrIle 79
Db 145 AGGACGGAACCGACGCTCCGATGCTACCTCTTCTCTGTCATGCTTGGCAGCCATT 204
Qy 80 AspLeuAlaAaSerValSerThrValProLysLeuLeuAlaIlePheTyrCysGlyAla 99
Db 205 GACCTGGCCTTATCCACATCCACCATGCTTAAGATCTTTCGCCCTTTCTGGTTGATGCC 264
Qy 100 GlyHisIleSerAlaSerAlaCysLeuAlaHisMetPhePheIleHisAlaPheCysMet 119
Db 265 CGAGAGATTAGCTTTGAGGCTGCTTACCAGATGTTCTTATTCATGCCCTCTCAGCC 324
Qy 120 MetGluSerThrValLeuLeuAlaMetAlaPheAspArgTyrValAlaIleCysHisPro 139
Db 325 ATTGAATCCACCATCTCTGCCATGCTTGGCCGTTATGTCGTCATCTGCCACCCA 384
Qy 140 LeuArgTyrAlaThrIleLeuThrAspThrIleIleAlaHisIleGlyValAlaVal 159
Db 385 CTGGCCCATGCTGAGTCTCAACATACAGTACAGCCAGATGGCATCTGGCTGTG 444
Qy 160 ValArgGlySerLeuLeuMetLeuProCysProPhePheIleGlyArgLeuAsnPheCys 179
Db 445 GTCCGCGGATCCCTCTTTTTCCTGCTGCTGCTGATCAAGCGCTGGCCTTCTGC 504
Qy 180 GlnSerHisValIleLeuHisThrTyrCysGluHisMetAlaValValLysLeuAlaCys 199
Db 505 CACTCCCATGCTCTCCGCTCTCTATTTGTCACAGGATGTAATGAAGTGGCCTAT 564
Qy 200 GlyAspThrArgProAsnArgValTyrGlyLeuThrAlaAlaLeuLeuValIleGlyVal 219
Db 565 GCAGACATTTGCCCAATGTGTATAGTCTTACTGCGCATCTCTGCTGTATGGGCGTG 624
Qy 220 AspLeuPheCysIleGlyLeuSerTyrAlaLeuIleAlaGlnAlaValLeuArgLeuSer 239
Db 625 GACGTAATGTTTCATCTCTCTGCTCTTCTGATAATACGAACGGTCTCTGCACTGCCT 684
Qy 240 SerHisGluAlaAaArgSerLysAlaLeuGlyThrCysGlySerHisValCysValIleLeu 259
Db 685 TCCAAGTCAGACGGCGGCGCAAGGCTTTTGGAAACCTGTGTGCACATTTGGTGTGATCTC 744

; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Hepier, William
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
; FILE REFERENCE: 210121.427C21
; CURRENT APPLICATION NUMBER: US/09/685.166A
; CURRENT FILING DATE: 2000-10-10
; NUMBER OF SEQ ID NOS: 898
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 526
; LENGTH: 963
; TYPE: DNA
; ORGANISM: Homo sapiens
US-08-465-980-1

Alignment Scores:
Pred. No.: 2.24e-91 Length: 1474
Score: 856.00 Matches: 160
Percent Similarity: 70.51% Conservative: 48
Best Local Similarity: 54.24% Mismatches: 87
Query Match: 49.83% Indels: 0
DB: 1 Gaps: 0

US-10-081-775-2 (1-329) x US-08-465-980-1 (1-1474)

Qy 24 LeuLeuGlyIleProGlyLeuGlnPheHisLeuTyrLeuSerLeuProValCysGly 43
Db 310 CTTATTGTATCCAGGATTAGAAAGCCCATTTCTGGTGGTCTCCCTCTCTTTC 369
```

```
QY 44 LeuGlyThrAlaThrIleValGlyAsnIleThrIleLeuValValAlaThrGluPro 63
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 370 ATGTATGTAGTGGCAATGTTGAAACTGCATCGTGTCTTCACTAAGGACGGAACGC 429
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 64 ValLeuHisGlyProValTyrLeuPheLeuCysMetLeuSerThrIleAspLeuAla 83
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 430 AGCCTGCACGCTCCGATGATACCTTCTCTGATGCTGCAGCATTGACCTGGCCTTA 489
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 84 SerValSerThrValProLysLeuAlaIlePheTyrCysGlyAlaGlyHisIleSer 103
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 490 TCACATCCACCATGCTAAGATCTTGGCCCTTTCTGTTTGTATTCGCGAGATTAGC 549
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 104 AlaSerAlaCysLeuAlaHisMetPheIleHisAlaPheCysMetMetGluSerThr 123
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 124 ValLeuLeuAlaMetAlaPheAspArgTyrValAlaIleCysHisProLeuArgTyrAla 143
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 610 ATCCTGCTGGCCATGGCCTTTGACCGTTATGTGGCCATCTGCCACCTGGCCATGCT 669
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 144 ThrIleLeuThrAspThrIleLeuAlaHisIleGlyValAlaValValArgGlySer 163
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 670 GCAGTGCTCAACAATACAGTAAACCCAGATGGCATCTGGCTGTGGTCCGCGATCC 729
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 164 LeuLeuMetLeuProCysProPhePheIleGlyArgLeuAsnPheCysGlnSerHisVal 183
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 730 CTCCTTTTTCCTCCACTGCCTCTGTGATCAAGCGCTGGCCCTTCTGCCACTCCAAATGTC 789
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 184 IleLeuHisThrTyrCysGluHisMetAlaValValLysLeuAlaCysGlyAspThrArg 203
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 790 CTCCTGCACCTCTATGTGTCCACGAGATGTAATCAAGTTGGCCCTATGCACACACTTTG 849
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 204 ProAsnArgValTyrGlyLeuThrAlaLeuLeuValIleGlyValAspLeuPheCys 223
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 850 CCCAATGGTATATGGTCTTACTGCCATTCGCTGGTCAATGGCGGTGAATGTTTC 909
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 224 IleGlyLeuSerTyrAlaLeuIleAlaGlnAlaValLeuArgLeuSerSerHisGluAla 243
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 910 ATCTCTTGCTCTATTTCTGATAATAGCAAGGTTCTGCAACTGCTTCAAGTCAGAG 969
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 244 ArgSerLysAlaLeuGlyThrCysGlySerHisValCysValIleLeuIleSerTyrThr 263
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 970 CGGCGCAAGGCTTTGGAACTGTGTGCACATTTGGTGTGTACTCGCCTCTATGTG 1029
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 264 ProAlaLeuPheSerPhePheThrHisArgPheGlyHisValProValHisIleHis 283
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 1030 CCATTATGGCCTCTCAGTTGTACACCGCTTTGGAAACAGCCCTTCATCCATTCATGCT 1089
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 284 IleLeuLeuAlaAsnValTyrLeuLeuLeuProProAlaLeuAsnProValValTyrGly 303
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 1090 GTTGTATGGGTGACATCTACTCTGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1149
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 304 ValLysThrLysGlnIleArgLysArgValValArgValPheGln 318
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 1150 GCCAAACCAACAGATCAGAACCGGCTGTGCTATGTTCAAG 1194
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
RESULT 6
US-09-053-303-1
; Sequence 1, Application US/09053303
; GENERAL INFORMATION:
; APPLICANT: Soppet, Daniel R.
; APPLICANT: Li, Yi
; APPLICANT: Li, Yi
; APPLICANT: Rosen, Craig A.
; APPLICANT: Ruben, Steven M.
; TITLE OF INVENTION: HUMAN G-PROTEIN RECEPTOR HPAJ70
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN, CECCHI,
; ADDRESSEE: STUART & OLSTEIN
; STREET: 6 Becker Farm Road
; CITY: Roseland
; STATE: New Jersey
```

```
; COUNTRY: USA
; ZIP: 07068
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/053,303
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/465,980
; FILING DATE: 06-JUN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Ferraro, Gregory D.
; REGISTRATION NUMBER: 36,134
; REFERENCE/DOCKET NUMBER: 325800-446
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-994-1700
; TELEFAX: 201-994-1744
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1474 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 274..1233
; US-09-053-303-1
Alignment Scores:
Pred. No.: 2,24e-91 Length: 1474
Score: 856.00 Matches: 160
Percent Similarity: 70.51% Conservative: 48
Best Local Similarity: 54.24% Mismatches: 87
Query Match: 49.83% Indels: 0
DB: 2 Gaps: 0
US-10-081-775-2 (1-329) x US-09-053-303-1 (1-1474)
QY 24 LeuLeuGlyIleProGlyLeuGluGlnPheHisLeuTrpLeuSerLeuProValCysGly 43
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 310 CTTATGGTATCCAGGATTAGAGAAAGCCCATTTCTGGTGTGGCTTCCCTCTCTTCC 369
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 44 LeuGlyThrAlaThrIleValGlyAsnIleThrIleLeuValValAlaThrGluPro 63
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 370 ATGTATGTAGTGGCAATGTTGGAACCTGTCATCGTGTCTTTCATCGTAAGGACGGAACGC 429
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 64 ValLeuHisLysProValTyrLeuPheLeuCysMetLeuSerThrIleAspLeuAla 83
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 430 AGCCTGCACGCTCCGATGATACCTTCTCTGATGCTTGCAGCATTTGACCTGGCCTTA 489
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 84 SerValSerThrValProLysLeuAlaIlePheTyrCysGlyAlaGlyHisIleSer 103
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 490 TCACATCCACCATGCTAAGATCTTGGCCCTTTCTGTTTGTATTCGCGAGATTAGC 549
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 104 AlaSerAlaCysLeuAlaHisMetPheIleHisAlaPheCysMetMetGluSerThr 123
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 550 ATTGAGGCCCTGCTTACCAGATGTTCTTATTATTCATGCCCTCTCAGCCATTGAATCCACC 609
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 124 ValLeuLeuAlaMetAlaPheAspArgTyrValAlaIleCysHisProLeuArgTyrAla 143
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 610 ATCCTGCTGGCCATGGCCTTTGACCGTTATGTGGCCATCTGCCACCTGGCCATGCT 669
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 144 ThrIleLeuThrAspThrIleLeuAlaHisIleGlyValAlaValValArgGlySer 163
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 670 GCAGTGCTCAACAATACAGTAAACCCAGATGGCATCTGGCTGTGGTCCGCGATCC 729
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 164 LeuLeuMetLeuProCysProPhePheIleGlyArgLeuAsnPheCysGlnSerHisVal 183
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
; COUNTRY: USA
; ZIP: 07068
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/053,303
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/465,980
; FILING DATE: 06-JUN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Ferraro, Gregory D.
; REGISTRATION NUMBER: 36,134
; REFERENCE/DOCKET NUMBER: 325800-446
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-994-1700
; TELEFAX: 201-994-1744
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1474 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 274..1233
; US-09-053-303-1
```


Qy 304 VallysthrLysGlnileArgLysArgValValArgValPheGln 318
 Db 1150 GCCAAACCAAACAGATCAGAACACGGGTGCTGGGTATGTTCAAG 1194

RESULT 8

```

PCT-US95-07093-1
; Sequence 1, Application PC/TUS9507093
; GENERAL INFORMATION:
; APPLICANT: Soppet, Daniel R.
; APPLICANT: Li, Yi
; APPLICANT: Rosen, Craig A.
; APPLICANT: Ruben, Steven M.
; TITLE OF INVENTION: HUMAN G-PROTEIN RECEPTOR HPRAJ70
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN, CECCHI,
; ADDRESSEE: STUART & OLSTEIN
; STREET: 6 Becker Farm Road
; CITY: Roseland
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07068
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/07093
; FILING DATE:

```

```

CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/465,980
FILING DATE: 06-JUN-1995
ATTORNEY/AGENT INFORMATION:
NAME: Ferraro, Gregory D.
REGISTRATION NUMBER: 36,7134
REFERENCE/DOCKET NUMBER: 325800-446
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201-994-1700
TELEFAX: 201-994-1744
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1474 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 274..1233
PCT-US95-07093-1

```

```

Alignment Scores:
  Cred. No.:      2.34e-91
  Score:          856.00
  Percent Similarity: 70.51%
  Local Similarity: 54.24%
  Query Match:    49.83%
  PB:              5
  Gaps:            0
  Indels:          0
  MismatchChes:   87
  Conservative:    48
  Matches:         160
  Length:         1474

```

US-10-081-775-2 (1-329) x PCT-US95-07093-1 (1-1474)

24	LeuLeuGlyIleProGlyLeuGluGlnPheHisLeuTriLeuSerLeuProValCysGly	43
310	CTTATTGGTATCCAGGATTAGAGAAAGCCATTCTGGGTGGCTTCCCTCTTTC	369
44	LeuGlyThrIleValGlyAsnIleThrIleuValValValIalaThrGluPro	63
370	ATGTATCTAGTGCATGTGTGAAACTGCATCGTGGCTTTCATGTAGGACGGAACGC	429
64	ValLeuHisLysProValTyrLeuPheLeuCysMetLeuSerThrIleAspLeuIala	83

Db	430	AGCTGCAGCGTCCGATGTACCTCTTTCTCGATGCTTGCGAGCAATGACCTGGCCTTA	489
Qy	84	SerValSerThrValProLysLeuLeuAlaIlePheTrpCysGlyAlaGlyHisSer	103
Db	490	TCACATCCACATGCGCTTAAGATCCTTGCCCTTTCTGGTTTGATTCGCCGAGATTAGC	549
Qy	104	AlaSerAlaCysLeuAlaHisMetPhePheIleHisAlaPheCysMetMetGluSerThr	123
Db	550	ATTGAGGCGCTGCTTACCAGATGTTCTTTATTATGCGCTCTCAGGCCATTGAATCCACC	609
Qy	124	ValLeuLeuAlaMetalPheAspArgTyrValAlaIleCysHisProLeuArgTyrAla	143
Db	610	ATCTGTGTGCCATGGCGCTTTGACGTTATGTGGCCATCTGCCACCCACTGGCGCATGCT	669
Qy	144	ThrIleLeuThrAspThrIleIleAlaHisIleGlyValAlaAlaValAlaArgGlySer	163
Db	670	GCAGTGTCAACATACAGTAACAGGCCACGATTGGCATCGTGGCTGTGGTCCGCGGATCC	729
Qy	164	LeuLeuMetLeuProCysProPhePheIleGlyArgLeuAsnPheCysGlnSerHisVal	183
Db	730	CTCTTTTTTCCCACATGCTCTGCTGATCAACGGCGCTGGCCCTCTGCCACTCCAAATGTC	789
Qy	184	IleLeuHisThrTyrCysGluHisMetAlaValValbysLeuAlaCysGlyAspThrArg	203
Db	790	CTTCGCACTCCTATTGTGTCCACGAGATGTAATGAATGGCCTATGCGACACACTTTG	849
Qy	204	ProAsnArgValTyrGlyLeuThrAlaAlaLeuLeuValIleGlyValAlaAspLeuPheCys	223
Db	850	CCCAATGTGTATGTGCTTACTGCCATTTCTGTGTCATGTGGCGTGGACGTAATGTTTC	909
Qy	224	IleGlyLeuSerTyrAlaLeuIleAlaGlnAlaValLeuArgLeuSerHisGluAla	243
Db	910	ATCTCTTGTCCTATTTCGTATAATACGACGGTTCGCAACTGCCCTTCCAAGTCAGAG	969
Qy	244	ArgSerIysAlaLeuGlyThrCysGlySerHisValCysValIleLeuIleSerTyrThr	263
Db	970	CGGGCCCAAGCGCTTTGGAACTGTGTGTACACATTTGGTGTGTACTCGCCTTCTATGTG	1029
Qy	264	ProAlaLeuPheSerPhePheThrHisArgPheGlyHisValProValHisIleHis	283
Db	1030	CCACTTATTGGCCCTCAGTTGTGTACACCGCTTTGGAAACAGCCCTTCATCCCATTTGTCGT	1089
Qy	284	IleLeuLeuAlaAsnValTyrLeuLeuLeuProProAlaLeuAsnProValValTyrGly	303
Db	1090	GTTGTATGGGTGACATCTACCTGCTGCTTCCTCTGTCATCAATCCCATCATCTAATGTT	1149
Qy	304	ValLysThrIysGlnIleArgIysArgValValArgValPheGln	318
Db	1150	GCCAAACCAACACAGATCAGAACACGGGTGCTGGCTATGTTCAAG	1194

RESULT 9

```

US-08-748-506-5
; Sequence 5, Application US/08748506
; Patent No. 6159707
; GENERAL INFORMATION:
; APPLICANT: Ronnett et al.
; TITLE OF INVENTION: NOVEL SPERM RECEPTORS
; NUMBER OF SEQUENCES: 31
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Leydig, Voit & Mayer, Ltd.
; STREET: Two Prudential Plaza, Suite 4900
; CITY: Chicago
; STATE: IL
; COUNTRY: US
; ZIP: 60601-6780
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/748,506
; FILING DATE: 08-NOV-1996

```

QY	268	serPhePheThrHisargPheGlyHisValProValHisIleHisIleLeuAla	267
Db	790	ACCTATTGAGGCCCAAGCTAGCCAC-----TCACCAGAAATGGACAAATTTCTTGCC	843
QY	288	AsnValTyrLeuLeuLeuProProAlaLeuAsnProValValTyrGlyValValYsThrLys	307
Db	844	CTCTCTCACAGTAGTGCATCATCCATGCTGAACCTATCATCTATATAGTTTAAAGGAACAAG	903
QY	308	GlnIleargLysArgValValArgValPheGlnSerGlyGlnGlyMetGlyIleLys	326
Db	904	GAAGTCAAGGCAGCAGCTAGAGAAGACT-----CTGGGCCTGAAA	942
RESULT 10			
US-08-748-506-7			
; Sequence 7, Application US/08748506			
; Patent No. 6153707			
; GENERAL INFORMATION:			
; APPLICANT: Ronnett et al.			
; TITLE OF INVENTION: NOVEL SPERM RECEPTORS			
; NUMBER OF SEQUENCES: 31			
; CORRESPONDENCE ADDRESS:			
; ADDRESSEE: Leydig, Voit & Mayer, Ltd.			
; STREET: Two Prudential Plaza, Suite 4900			
; CITY: Chicago			
; STATE: IL			
; COUNTRY: US			
; ZIP: 60601-6780			
; COMPUTER READABLE FORM:			
; MEDIUM TYPE: Floppy disk			
; COMPUTER: IBM PC compatible			
; OPERATING SYSTEM: PC-DOS/MS-DOS			
; SOFTWARE: PatentIn Release #1.0, Version #1.25			
; CURRENT APPLICATION DATA:			
; APPLICATION NUMBER: US/08/748,506			
; FILING DATE: 08-NOV-1996			
; CLASSIFICATION: 435			
; PRIOR APPLICATION DATA:			
; APPLICATION NUMBER: US 60/033,751			
; FILING DATE: 09-NOV-1995			
; CLASSIFICATION: 435			
; ATTORNEY/AGENT INFORMATION:			
; REFERENCE/DOCKET NUMBER: 74940			
; TELECOMMUNICATION INFORMATION:			
; TELEPHONE: 312-616-5600			
; TELEFAX: 312-616-5700			
; INFORMATION FOR SEQ ID NO: 7:			
; SEQUENCE CHARACTERISTICS:			
; LENGTH: 966 base pairs			
; TYPE: nucleic acid			
; STRANDEDNESS: double			
; TOPOLOGY: linear			
; MOLECULE TYPE: DNA (genomic)			
US-08-748-506-7			
Alignment Scores:			
Pred. No.: 1.81e-40 Length: 966			
Score: 423.50 Matches: 105			
Percent Similarity: 52.00% Conservative: 51			
Best Local Similarity: 35.00% Mismatches: 133			
Query Match: 24.65% Indels: 11			
DB: 3 Gaps: 4			
US-10-081-775-2 (1-329) x US-08-748-506-7 (1-966)			
QY	27	IleProGlyLeuGluGlnPheHisLeuThrLeuSerLeuProValCysGlyLeuGlyThr	46
Db	76	GTCCCTGGA--GAATGCTTCCTCTGTTCCACCTCATCTCTCATGTCTTA-----	126
QY	47	AlaThrIleValGlyAsnIleThrIleLeuValValAlaThrGluProValLeuHis	66
Db	127	GTATCACTAACAGGAATACTCTCATAGCCCTTGCTTTGTACCAAGTCATCTCTACAC	196
QY	67	IysProValTyrLeuPheLeuCvsMetLeuSerThrIleAspLeuAlaIaSerValSer	86


```

Db 187 ACCCCCATGTACTCTCTCTGGCAACTTGTCTCTCCCTGAGATTGGCTATATCTGCTCT 246
Qy 87 ThrValProLysLeuLeuAlaIlePheTrpCysGlyAlaGlyHisIleSerAlaSerAla 106
Db 247 GTCATACCAAGATCTCTCTGAGCCCTTGTGAGTGGCCGAGGATCTCTGGGAGGCT 306
Qy 107 CysLeuAlaHisMetPhePheIleHisAlaPheCysMetMetGluSerThrValLeuLeu 126
Db 307 TGTGCTCACAGATCTCTCTCATATCTTTGTATTAAGTGTGCTGCTATTTGGCA 366
Qy 127 AlaMetAlaPheAspArgTyrValAlaIleCysHisProLeuArgTyrAlaThrIleLeu 146
Db 367 GCCATGGCTTGTGACCGCTATATGCTATATGCTTCCCTCCCTCCCTATGCAACCGAATG 426
Qy 147 ThrAspThrIleIleAlaHisIleGlyValAlaAlaValAlaValArgGlySerLeuLeuMet 166
Db 427 AGTCGTGGGTATGCTGCTATTTGGCAATGCTCTCATGGTGTGATGGATGATAGT 486
Qy 167 LeuProCysProPhePheIleGlyArgLeuAsnPheCysGlnSerHisValIleLeuHis 186
Db 487 CTGGGACAGACCAATTTATTTCTCTTGAACCTTCTGTGGACCTGTGAGATGACAC 546
Qy 187 ThrTyrCysGluHisMetAlaValValLysLeuAlaCysGlyAspThrArgProAsnArg 206
Db 547 TTCTTCTGTGACCTTCCACCTCTCTGTCGACCTTGTGCTGTGATATACATCCCAATGAG 606
Qy 207 ValTyrGlyLeuThrAlaAlaLeuValIleGlyValAlaAspLeuPheCysIleGlyLeu 226
Db 607 GCTGCCATCTTTGGGACGAGTGTCTGCAATTTAGTCCATTTTACTGATCTTCT 666
Qy 227 SerTyrAlaLeuIleAlaGlnAlaValLeuArgLeuSerHisGluAlaArgSerLys 246
Db 667 TCCTATGTCAGAAATCTCGTGTGAGTGTGCTGTGATGCTTCCCTGAGGGCGCAATAA 726
Qy 247 AlaLeuGlyThrCysGlySerHisValCysValIleLeuIleSerTyrThrProAlaLeu 266
Db 727 GCTCTCTCTACCTGTCTCATCTCACTTCTGTAGTCACACTTCTATGGCTCAACATCT 786
Qy 267 PheSerPhePheThrHisArgPheGlyHisValProValHisIleHisIleLeuLeu 286
Db 787 GCCACCTTTGAGGTCAAGTGTAGCCATCTAGCCATCCACGAGAGTG-----GACAAACTCTTG 840
Qy 287 AlaAsnValTyrLeuLeuProAlaLeuAsnProValValTyrValTyrGlyValLysThr 306
Db 841 GCCCTCTCTATACATCAGTGACATCCATGCTGTGATCCATCCATCTACAGCTTAAGGAAC 900
Qy 307 LysGlnIleArgLysArgValValArgValPheGlnSerGlyGlnGlyMetGlyIleLys 326
Db 901 AAGGAAGTAAGGGTGTCACTGAGAAGAACT-----CTGGGCTGAAG 942

```

RESULT 11

```

US-08-748-506-6
; Sequence 6, Application US/08748506
; Patent No. 6159707
; GENERAL INFORMATION:
; APPLICANT: Ronnett et al.
; TITLE OF INVENTION: NOVEL SPERM RECEPTORS
; NUMBER OF SEQUENCES: 31
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Leydig, Voit & Mayer, Ltd.
; STREET: Two Prudential Plaza, Suite 4900
; CITY: Chicago
; STATE: IL
; COUNTRY: US
; ZIP: 60601-6780
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/748,506

```

```

; FILING DATE: 08-NOV-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/033,751
; FILING DATE: 09-NOV-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; REFERENCE/DOCKET NUMBER: 74940
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-616-5600
; TELEFAX: 312-616-5700
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 966 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-748-506-6

Alignment Scores:
Pred. No.: 1,22e-39 Length: 966
Score: 416.50 Matches: 102
Percent Similarity: 50.17% Conservative: 48
Best Local Similarity: 34.11% Mismatches: 144
Query Match: 24.24% Indels: 5
DB: 3 Gaps: 3

```

```

US-10-081-775-2 (1-329) x US-08-748-506-6 (1-966)
Qy 27 IleProGlyLeuGluInPheHisIleThrLeuSerLeuProValCysGlyLeuGlyThr 46
Db 76 GTCCTTGA---GAATGCTCTCTCTCTGTCACCCCTCATCTCTCTCATGTTCTTA----- 126
Qy 47 AlaThrIleValGlyAsnIleThrIleLeuValValAlaThrGluProValLeuHis 66
Db 127 GTATCATACAGAAATGCTCTCATAGCCCTTGTGTTGTACAGTCCATCTCTACAC 186
Qy 67 LysProValTyrLeuPheLeuCysMetLeuSerThrIleAspLeuAlaAlaSerValSer 86
Db 187 ACCCCCATGTACTCTTCTTGGCCAACTGTCTCTCTGGAGATGGCTATATCTGCTCT 246
Qy 87 ThrValProLysLeuLeuAlaIlePheTrpCysGlyAlaGlyHisIleSerAlaSerAla 106
Db 247 GTCATACCAAGATGCTGTCAGAGCCTTGTGAGTGGCCGAGAGATCTCTAGGAGGGA 306
Qy 107 CysLeuAlaHisMetPhePheIleHisAlaPheCysMetMetGluSerThrValLeuLeu 126
Db 307 TGTGCCACAGATGTTTCTTTCACATTTTGGCATAACTGAGTGTGCTGCTTATGGCA 366
Qy 127 AlaMetAlaPheAspArgTyrValAlaIleCysHisProLeuArgTyrAlaThrIleLeu 146
Db 367 GCCATGGCTTGTGACCGCTATATGCTTCCCTCCCTATGCAACCGAATG 426
Qy 147 ThrAspThrIleIleAlaHisIleGlyValAlaAlaValAlaValArgGlySerLeuLeuMet 166
Db 427 AGTCGTGAGTATGTGCCCAATTTGGCAATTTGTCAGTGGGAAATGGGATGATAGT 486
Qy 167 LeuProCysProPhePheIleGlyArgLeuAsnPheCysGlnSerHisValIleLeuHis 186
Db 487 CTGGGACAGACCAATTTATTTCTTGAACCTTCTGTGGACCTGTGAGATGACAC 546
Qy 187 ThrTyrCysGluHisMetAlaValValLysLeuAlaCysGlyAspThrArgProAsnArg 206
Db 547 TTCTTCTGTGACCTTCCACCTCTCTGTCGACCTTGTGCTGTGATATACATCCCAATGAG 606
Qy 207 ValTyrGlyLeuThrAlaAlaLeuValIleGlyValAlaAspLeuPheCysIleGlyLeu 226
Db 607 GCTGCCATCTTTGGGACGAGTGTCTGCAATTTAGTCCATTTTACTGATCTTCT 666
Qy 227 SerTyrAlaLeuIleAlaGlnAlaValLeuArgLeuSerHisGluAlaArgSerLys 246
Db 667 TCTTATGTCAGAAATCTCGTGTGAGTGTGCTGATGCTTCTTCTGAGGCGCAATAA 726

```

```
QY 247 AlaLeuGlyThrCysGlySerHisValCysValIleLeuLeuSerThrProAlaLeu 266
Db 727 GCTCTTCCACCTGCTCCTCACACCTACTTGTAGTCACACTCTTTTATGGCTCTGTCTCC 786
QY 267 PheSerPhePheThrHisArgPheGlyHisValProValHisIleHisIleLeu 286
Db 787 TTTACCTATTGGAGCCCAAGTCTAGCCAC-----TCACAGGAATGGACAACTCTTG 840
QY 287 AlaAsnValTyrLeuLeuLeuProAlaLeuAsnProValValTyrGlyValIleThr 306
Db 841 GCCCTCTTACACAGCAGTGCATCCATGCTGAACCTCATCTACAGCTTAAGGAAC 900
QY 307 LysGlnIleArgLysArgValValArgValPheGlnSerGlyGlnGlyMetGlyLe 325
Db 901 AAGGAAGTCAAGGCAGCAGTCTGAGAAGAACTCTCGACCTGAAAAAATATATGTCATT 957

RESULT 12
US-08-748-506-8
; Sequence 8, Application US/08748506
; Patent No. 6159707
; GENERAL INFORMATION:
; APPLICANT: Ronnett et al.
; TITLE OF INVENTION: NOVEL SPERM RECEPTORS
; NUMBER OF SEQUENCES: 31
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Leydig, Voit & Mayer, Ltd.
; STREET: Two Prudential Plaza, Suite 4900
; CITY: Chicago
; STATE: IL
; COUNTRY: US
; ZIP: 60601-6780
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; FILING DATE: 08-NOV-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/033,751
; FILING DATE: 09-NOV-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; REFERENCE/DOCKET NUMBER: 74940
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-616-5600
; TELEFAX: 312-616-5700
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 966 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-748-506-8

Alignment Scores:
Pred. No.: 6.27e-39 Length: 966
Score: 410.50 Matches: 99
Percent Similarity: 50.6% Conservatives: 48
Best Local Similarity: 34.14% Mismatches: 138
Query Match: 23.89% Indels: 5
DB: 3 Gaps: 3

US-10-081-775-2 (1-329) x US-08-748-506-8 (1-966)
QY 27 IleProGlyLeuGlnPheHisIleTyrLeuSerLeuProValCysGlyLeuGlyThr 46
Db 76 GTCCCTGGG---GAATGCTTCTCCTGTTTCAACCTCATCTTCTCATGTTCTTCTTA----- 126
```

```
QY 47 AlaThrIleValGlyAsnIleThrIleLeuValValAlaThrGluProValLeuHis 66
Db 127 GTATCACTAACAGAAATGCTCTCATAGCCCTTCTATTTGTACCACTCCATCTCTACAC 186
QY 67 LysProValTyrLeuPheLeuCysMetLeuSerThrIleAspLeuAlaIleSerValSer 86
Db 187 ACCCCCATGTACTCTTTCTGGCCAACTGTCTCTCTGGAGATTGGCTATCTACTTGTCT 246
QY 87 ThrValProLysLeuLeuAlaIlePheTrpCysGlyAlaGlyHisIleSerAlaSerAla 106
Db 247 GTCATACCCCAAGATGCTGACAGAGTCTTGTAGTGGCCGAGAGATCTTTCAGTGGGA 306
QY 107 CysLeuAlaHisMetPhePheIleHisAlaPheCysMetMetGluSerThrValLeuLeu 126
Db 307 TGTGCCACACAGATGTTTCTTCATATCTTCTGGTATAAAGTGTGCTGCTATTGGCA 366
QY 127 AlaMetAlaPheAspArgTyrValAlaIleCysHisProLeuArgTyrAlaThrIleLeu 146
Db 367 GCCATGGCTTTGACCGCTATATGGCTATATGTTCCCACTCCACTATGCAACCCGAATG 426
QY 147 ThrAspThrIleIleAlaHisIleGlyValAlaValAlaValArgGlySerLeuMet 166
Db 427 AGTCGTAGGTATGTGCCCACTTGGCAATGTTTTCATGGGTGATGGATGCAATAGTAGGT 486
QY 167 LeuProCysProPhePheIleGlyArgLeuAsnPheCysGlnSerHisValIleLeuHis 186
Db 487 CTGGGACAGACCAATTTTATTTCTCTTGAACCTCTGTGCACCTGTGAGATAGACCCAC 546
QY 187 ThrTyrCysGluHisMetAlaValLysLeuAlaCysGlyAspThrArgProAsnArg 206
Db 547 TTCTTCTGTGATCTTCCACCTCTCTCGGCATCTGCTGTGTGTATACATCCCAATTCAG 606
QY 207 ValTyrGlyLeuThrAlaAlaLeuLeuValIleGlyValAspLeuPheCysIleGlyLeu 226
Db 607 GCTGCCATCTTTGTGGTAGTGTCTCTGCAATATCTAGCCCTTTTTCCTGATCATTTAT 666
QY 227 SerTyrAlaLeuIleAlaGlnAlaValLeuArgLeuSerSerHisGlnAlaArgSerLys 246
Db 667 TCTTATGTCAAGATCTCGTTGAGTGTGCTGTGATGCCCTTCACTGAGGGCGGCACAAA 726
QY 247 AlaLeuGlyThrCysGlySerHisValCysValIleLeuIleSerTyrThrProAlaLeu 266
Db 727 GCCCTTTCAACCTGTCTCCACCTACTTGTAGTCACACTCTTTTATGGCTCAGGATCT 786
QY 267 PheSerPhePheThrHisArgPheGlyHisValProValHisIleHisIleLeuLeu 286
Db 787 GTTACCTATTGAGCCCTAAGTCTAGCCAC-----TCACAGGAATGGACAACTCTTG 840
QY 287 AlaAsnValTyrLeuLeuLeuProAlaLeuAsnProValValTyrGlyValIleThr 306
Db 841 GCCCTCTTCTACACAGCAGTGCATCCATGTTGAACCTCATCTATCATATATGTTTAAAG 900
QY 307 LysGlnIleArgLysArgValValArgVal 316
Db 901 AAGGATGTCAAGGCAGCAGTCTGAGAAGAAAT 930

RESULT 13
US-09-016-434-1413
; Sequence 1413, Application US/09016434
; Patent No. 6500938
; GENERAL INFORMATION:
; APPLICANT: Janice Au-Young
; APPLICANT: Jeffrey J. Seilhamer
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING
; TITLE OF INVENTION: PATHWAY GENE EXPRESSION
; NUMBER OF SEQUENCES: 1490
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
```

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/016,434
FILING DATE: HEREWITH
CLASSIFICATION:
PRIORITY APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Zeller, Karen J.
REGISTRATION NUMBER: 37,071
REFERENCE/DOCKET NUMBER: PA-0002 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 855-0555
TELEFAX: (650) 845-4166
INFORMATION FOR SEQ ID NO: 1413:
SEQUENCE CHARACTERISTICS:
LENGTH: 1282 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: GENBANK
CLONE: G516319
US-09-016-434-1413

Alignment Scores:
Pred. No.: 1,12e-38 Length: 1282
Score: 410.00 Matches: 97
Percent Similarity: 51.00% Conservative: 56
Best Local Similarity: 32.33% Mismatches: 135
Query Match: 23.86% Indels: 12
DB: 4 Gaps: 4

US-10-081-775-2 (1-329) x US-09-016-434-1413 (1-1282)

```
QY 22 PhePheLeuLeuGlyLeu-----ProGlyLeuGluGln-----PheHisLeuTirp 36
Db 201 TTATCCTGCTGGGCTTGTGGAGGCGCCAGGCTGCAGCCAGTGTCTTGTCTCTTC 260
QY 37 LeuSerLeuProValCysGlyLeuGlyThraAlaThrIleValGlyAenIleThrIleLeu 56
Db 261 CTCTTTGCCTTACCTG-----GTCACGGTGCAGGGCAACCTCAGCATCTG 305
QY 57 ValValAlaThrGluProValLeuHisLysProValTyrLeuPheLeuCysMetLeu 76
Db 306 GCAGCTGCTGTGGAGGCGCAACTCCACCCCACTGACTTCTTCTTCCCACTG 365
QY 77 SerThrIleAspLeuAlaSerValSerThrValProLysLeuLeuAlaIlePheTirp 96
Db 366 TCAGTGTGTCATGTGGGTGCATCAGCGTCATCTGCTTCCATCAATGTGAGTGTCTCTG 425
QY 97 CysGlyAlaGlyHisIleSerAlaSerAlaCysLeuAlaHisMetPhePheIleHisAla 116
Db 426 TCCCGCAAGCGTGCAGTTCCTGTGGGGCGCTGCTTACCCAGCTCTTCTTCTTCCATCTG 485
QY 117 PheCysMetMetGluSerThrValLeuAlaMetAlaPheAspArgTyrValAlaIle 136
Db 486 TTCTGTGGAGTGCAGTCTTCTGCTGACCGCATGCGCTATGACCAATTCCTGGGCATC 545
QY 137 CysHisProLeuArgTyrAlaThrIleLeuThrAspThrIleIleAlaHisIleGlyVal 156
Db 546 TGGCGGCGCCCTACCTACAGCCCGCATGAGTCAGACAGTCCAGAGGATGTGGTGGCT 605
QY 157 AlaAlaValValArgGlySerLeuLeuMetLeuLeuProCysProPhePheIleGlyA-gLeu 176
Db 606 GCGTCTCTGGGCTGTGCTTTTACCAACGCACTGACCCACACATGCTGGGCCATGTCCAGCTC 665
```

```
QY 177 AsnPheCysGlnSerHisValIleLeuHisThrTyrCysGluHisMetAlaValValLys 196
Db 666 AACTTCTGTGGCCCAATGTGATCAATCACTTCTACTGTGACCTCCACAGCTCTTCCAG 725
QY 197 LeuAlaCysGlyAspThrArgProAsnArgValTyrGlyLeuThrAlaAlaLeuVal 216
Db 726 CTCTCTGCTCCAGCACCAACTCAATGAGCTGTGCTCTTTTGTGTGGTTTATAATG 785
QY 217 IleGlyValAspLeuPheCysIleGlyLeuSerTyrAlaLeuIleAlaGlnAlaValLeu 236
Db 786 GCAGGTACCCCATGCTCTCATTTCTATCTATATCCAGTGCAGCTGCAGTCCCTG 845
QY 237 ArgLeuSerSerHisGluAlaArgSerLysAlaLeuGlyThrCysGlySerHisValCys 256
Db 846 CGAATTCGCTCTGTAGAGGCGAGGAGAAAGCTTCTCCACATGTGGTCCACCTCACT 905
QY 257 ValIleLeuIleSerTyrThrProAlaLeuPheSerPhePheThrHisArgPheGlyHis 276
Db 906 GTGGTGGCATATCTATGTTGAGTATCTTAACTATATG-----CGACTGGGTTC 959
QY 277 HisValProValHisIleHisIleLeuLeuAlaAsnValTyrLeuLeuProProAla 296
Db 960 ACCAAGCTTTCAGACAGGATAAAGCTGTTGGAATTTTCAACACTGTCTCAATCCCATG 1019
QY 297 LeuAsnProValValTyrGlyValLysThrLysGlnIleArgLysArgValValArgVal 316
Db 1020 CTGAACCAATCATCTACAGCTTCAGAAACCTGATGTGAGAGTGCCATCTGGAGGATG 1079
RESULT 14
US-09-546-986A-5
; Sequence 5, Application US/09546986A
; Patent No. 6635741
; GENERAL INFORMATION:
; APPLICANT: Powers, Scott
; APPLICANT: Yang, Jianxin
; APPLICANT: Cutler, Gene
; APPLICANT: Tularik Inc.
; TITLE OF INVENTION: No. 6635741el G-Protein Coupled Receptors
; FILE REFERENCE: 018781-004720US
; CURRENT APPLICATION NUMBER: US/09/546,986A
; PRIOR FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 09/524,730
; PRIOR FILING DATE: 2000-03-14
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 1351
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (86)..(1108)
; OTHER INFORMATION: human breast cancer amplified G-protein coupled
; OTHER INFORMATION: receptor 3 (BCA-GPCR-3)
US-09-546-986A-5
```

Alignment Scores:
Pred. No.: 1,21e-38 Length: 1351
Score: 410.00 Matches: 101
Percent Similarity: 48.62% Conservative: 57
Best Local Similarity: 31.08% Mismatches: 131
Query Match: 23.86% Indels: 36
DB: 4 Gaps: 5

US-10-081-775-2 (1-329) x US-09-546-986A-5 (1-1351)

```
QY 6 GlyHisAsnMetGluSerProHisHisThrAspValAsp-----Pro 19
Db 116 GGTGGCTTTTGGCCCACTCCAGCAGCATAATGGAATAGCAATGTGAGTCTCTCCA 175
QY 20 SerValPhePheLeuGlyIleProGlyLeuGluGlnPheHisLeuTrpLeuSerLeu 39
Db 176 GAAGTCTTGTCTCTGGGCTTCTCCGACGACCCCTCCTAGAAACTGTCTCTTCTATA 235
```

; CURRENT FILING DATE: 2000-03-14
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 1351
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (86)..(1108)
; OTHER INFORMATION: human breast cancer amplified G-protein coupled
; OTHER INFORMATION: receptor 3 (BCA-GPCR-3)
US-09-524-730-5

Alignment Scores:
Pred. No.: 1.21e-38 Length: 1351
Score: 410.00 Matches: 101
Percent Similarity: 48.62% Conservative: 57
Best Local Similarity: 31.08% Mismatches: 131
Query Match: 23.86% Indels: 36
Gaps: 5
DB:

US-10-081-775-2 (1-329) x US-09-524-730-5 (1-1351)
QY 6 GlyHisAsnMetGluSerProHisHisThrAspValAsp-----Pro 19
DB 116 GGTGGCCTTTTGGCCCCCAGCATACATGGAATAGCAATGTGAGTTCTCCA 175
QY 20 SerValPhePheLeuGlyLeuProGlyLeuGluGlnPheHisLeuTrpLeuSerLeu 39
DB 176 CAAGCTCTTTGCTCTCTGGGCTTCTCCGACGACCCCTCCTAGAACTGTCTCTTCATA 235
QY 40 ProValCysGlyLeuGlyThrAlaThrIleValGlyAsnIleThrIleLeuValVal 59
DB 236 GTTGCTTTGAGTTTACATGATCGATCTGGCATATGTCATCATCTTGTCTCC 295
QY 60 AlaThrGluProValLeuHisLysProValTyrLeuPheLeuCysMetLeuSerThrIle 79
DB 296 CATACAGATGTGCACCTCCACACACCTATGCTTCTTTCTTTCGCAACCTCTCTCTG 355
QY 80 AspLeuAlaAlaSerValSerThrValProLysLeuLeuAlaIlePheTrpCysGlyAla 99
DB 356 GACATGAGCTTCACACGACGATTTGTCACAGCTCTGGGTAACTCTGGGACACAG 415
QY 100 GlyHisIleSerAlaSerAlaCysLeuAlaHisMetPhePheIleHisAlaPheCysMet 119
DB 416 AAACCATATAGCTATGAGGGTGTGTGTCAGTCTATATCTCCATTCCTGGGGGCA 475
QY 120 MetGluSerThrValLeuAlaMetAlaPheAspArgTyrValAlaIleCysHisPro 139
DB 476 ACCGAGTGTCTCTGCTGCGCCACCATGCTCTATGACCGCTACGCTGCCATCTGCAGGCCA 535
QY 140 LeuArgTyrAlaThrIleLeuThrAspThrIleIleAlaHisIleGlyValAla----- 157
DB 536 CTCCTATTACTGCTATGATCCACAGCTTTGCCCTTGGCTAGCTTTGGCTCTCTGG 595
QY 158 -----AlaValValArgGlySerLeuLeuMetLeuProCysProPhe 171
DB 596 CTGGGGGGTCTGACCCACGACGATGTGGCTCCAGCTCACCCTGCTC----- 643
QY 172 PheIleGlyArgLeuAsnPheCysGlnSerHisValIleLeuHisThrTyrCysGluHis 191
DB 644 -----CTACCGCTGTGTGGGAACAATTCATCGACCTGACCTCTTTTGGGAGATG 691
QY 192 MetAlaValValLysLeuAlaCysGlyAspThrArgProAsnArgValTyrGlyLeuThr 211
DB 692 CCCCTCATTTGCACTGGCTTGTGTGGATACAGCTCAATGAGATGAGATGATGACCTG 751
QY 212 AlaAlaLeuValIleGlyValAspLeuPheCysIleGlyLeuSerTyrAlaLeuIle 231
DB 752 GCCAGCTTTGTCTTTGTTGCTCTGCTGGGCTCTCTCTGCTCTCTTACGCCACATT 811
QY 232 AlaGlnAlaValLeuArgLeuSerSerHisGluAlaArgSerLysAlaLeuGlyThrCys 251

QY 40 ProValCysGlyLeuGlyThrAlaThrIleValGlyAsnIleThrIleLeuValValVal 59
DB 236 GTTGCTTTGAGTTTACATGATCGATCTGGGCAATGGCATCATCAITTTGTCTCC 295
QY 60 AlaThrGluProValLeuHisLysProValTyrLeuPheLeuCysMetLeuSerThrIle 79
DB 296 CATACAGATGTGCACCTCCACACACCTATGCTTCTTTGCAACCTCTCTCTCTG 355
QY 80 AspLeuAlaAlaSerValSerThrValProLysLeuLeuAlaIlePheTrpCysGlyAla 99
DB 356 GACATGAGCTTCACACGACGATTTGTCACAGCTCTGGTAACTCTGGGACACAG 415
QY 100 GlyHisIleSerAlaSerAlaCysLeuAlaHisMetPhePheIleHisAlaPheCysMet 119
DB 416 AAACCATATAGCTATGAGGGTGTGTGTCAGTCTATATCTCCATTCCTGGGGGCA 475
QY 120 MetGluSerThrValLeuAlaMetAlaPheAspArgTyrValAlaIleCysHisPro 139
DB 476 ACCGAGTGTCTCTGCTGCGCCACCATGCTCTATGACCGCTACGCTGCCATCTGCAGGCCA 535
QY 140 LeuArgTyrAlaThrIleLeuThrAspThrIleIleAlaHisIleGlyValAla----- 157
DB 536 CTCCTATTACTGCTATGATCCACAGCTTTGGCTTGGCTAGCTTTGGCTCTCTGG 595
QY 158 -----AlaValValArgGlySerLeuLeuMetLeuProCysProPhe 171
DB 596 CTGGGGGGTCTGACCCACGACGATGTGGCTCCAGCTCACCCTGCTC----- 643
QY 172 PheIleGlyArgLeuAsnPheCysGlnSerHisValIleLeuHisThrTyrCysGluHis 191
DB 644 -----CTACCGCTGTGTGGGAACAATTCATCGACCTGACCTCTTTTGGGAGATG 691
QY 192 MetAlaValValLysLeuAlaCysGlyAspThrArgProAsnArgValTyrGlyLeuThr 211
DB 692 CCCCTCATTTGCACTGGCTTGTGTGGATACAGCTCAATGAGATGAGATGATGACCTG 751
QY 212 AlaAlaLeuValIleGlyValAspLeuPheCysIleGlyLeuSerTyrAlaLeuIle 231
DB 752 GCCAGCTTTGTCTTTGTTGCTCTGCTGGGCTCTCTGCTCTTACGGCCACATT 811
QY 232 AlaGlnAlaValLeuArgLeuSerSerHisGluAlaArgSerLysAlaLeuGlyThrCys 251
DB 812 GCCCGGGCGGTGTGAAGATCAGTTCACAGAGAGGGGGAGGAGGACCTTCAACACCTGT 871
QY 252 GlySerHisValCysValIleLeuIleSerTyrThrProAlaLeuPheSerPhePhe--- 270
DB 872 TCTTCCACAGTGGCTGTGTGTCTCTGTTTACGGGAGCATCTCTCATGTATCTCCAG 931
QY 271 -----ThrHisArgPheGlyHisHisValProValHisIleIleLeu 285
DB 932 CCAGCCAAAGACACCTCCCATGACGAGGCAAG-----TTC 967
QY 286 LeuAlaAsnValTyrLeuLeuLeuProProAlaLeuAsnProValTyrGlyValLys 305
DB 968 ATAGCTCTGTTTACACCGTAGTCTCTGTTTACGGGAGCATCTTATTTACCCCTGAGG 1027
QY 306 ThrLysGlnIleArg 310
DB 1028 AACACGAGGTGAAG 1042

RESULT 15
US-09-524-730-5
; Sequence 5, Application US/09524730
; Patent No. 6638733
; GENERAL INFORMATION:
; APPLICANT: Powers, Scott
; APPLICANT: Yang, Jianxin
; APPLICANT: Cutler, Gene
; APPLICANT: Tularik Inc.
; TITLE OF INVENTION: No. 6638733el G-Protein Coupled Receptors
; FILE REFERENCE: 018781-004710US
; CURRENT APPLICATION NUMBER: US/09/524, 730

```

Db      812 GCGCGGCGGTGTGAAGATCAGGTACAGCAGAGGCGGAGAGGCATTCAACACCTGT 871
QY      252 GlySerHisValCysValIleLeuIleSerTyrThrProAlaLeuPheSerPhePhe--- 270
Db      872 TCCTCCACAGTGGCTGTGTCTCTGTTTACGGGAGCATCATCTTCATGTATCTCCAG 931
QY      271 -----ThrHisArgPheGlyHisHisValProValHisIleLeu 285
Db      932 CCAGCCAGAGCACCTCCCATGAGCAGGCAAG-----TTC 967
QY      286 LeuAlaAsnValTyrLeuLeuLeuProProAlaLeuAsnProValValTyrGlyVallys 305
Db      968 ATAGCTCTGTCTACACCGTAGTCACTCCTGGTTGAACCCACTTATTACACCCCTGAGG 1027
QY      306 ThrLysGlnIleArg 310
Db      1028 AACCGGAGGTGAAG 1042

```

Search completed: August 27, 2004, 20:57:49
Job time : 98 secs

This Page Blank (uspio)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: August 27, 2004, 18:28:44 ; Search time 3045 Seconds

(without alignments)
3226.488 Million cell updates/sec

Title: US-10-081-775-2

Perfect score: 1718

Sequence: 1 MSSTLGHNMESPHTDVPDPS.....RKRVRVRFSGQGMGIKASE 329

Scoring table:

BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-O=/cgn2_1/USPTO_spool/US10081775/runat_20082004_170213_11139/app_query.fasta_1.519
-DB=BST_QFMT=fastcap -SUFFIX=first -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10081775 @CGN 1 1 3437 @runat_20082004_170213_11139 -NCPUP=6 -ICPU=3
-NO.MMAP -LARGQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=40 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOF=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hcc:*
9: gb_est1:*
10: gb_est2:*
11: gb_hcc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vit:*
21: em_gss_fun:*
22: em_gss_man:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_phg:*
27: em_gss_vri:*
28: gb_gss1:*

29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	977.5	56.9	948	29	AY401474	Mus muscu
2	961.5	56.0	948	29	AY401472	AY401472 Homo sapi
3	904	52.6	936	29	AY401478	AY401478 Homo sapi
4	889.5	51.8	948	29	AY401473	AY401473 Pan trogl
5	889.5	51.8	957	29	AY414225	AY414225 Homo sapi
6	882.5	51.4	954	29	AY414227	AY414227 Mus muscu
7	880.5	51.3	936	29	AY401480	AY401480 Mus muscu
8	868	50.5	1964	11	AK036356	AK036356 Mus muscu
9	868	50.5	3410	11	AK028467	AK028467 Mus muscu
10	867.5	50.5	948	29	AY410600	AY410600 Mus muscu
11	834	48.5	948	29	AY410598	AY410598 Homo sapi
12	828.5	48.2	846	29	AY402272	AY402272 Mus muscu
13	827	48.1	948	29	AY410599	AY410599 Pan trogl
14	819	47.7	966	29	AY405569	AY405569 Mus muscu
15	802	46.7	1086	28	AF101706	AF101706 Mus muscu
16	794	46.2	822	29	CC500683	CC500683 CH240_338
17	788	45.9	966	29	AY405567	AY405567 Homo sapi
18	784	45.6	900	29	AY410597	AY410597 Mus muscu
19	769	44.8	900	29	AY410595	AY410595 Homo sapi
20	768.5	44.7	846	29	AY402270	AY402270 Homo sapi
21	762	44.4	900	29	AY410596	AY410596 Pan trogl
22	750	43.7	774	29	CC512357	CC512357 CH240_356
23	730.5	42.5	746	29	AY401479	AY401479 Pan trogl
24	730	42.5	948	29	AY402752	AY402752 Mus muscu
25	717.5	41.8	957	29	AY414226	AY414226 Pan trogl
26	713.5	41.5	939	29	AY402268	AY402268 Pan trogl
27	704.5	41.0	798	28	BH082934	BH082934 RPT-24-9
28	703.5	40.9	939	29	AY402267	AY402267 Homo sapi
29	696	40.5	963	29	AY401651	AY401651 Mus muscu
30	678.5	39.5	853	29	CC527411	CC527411 CH240_402
31	673.5	39.2	966	29	AY401649	AY401649 Homo sapi
32	671	39.1	864	29	CC480759	CC480759 CH240_308
33	669.5	39.0	820	28	BH046083	BH046083 RPT-24-2
34	663.5	38.6	987	29	AY402269	AY402269 Mus muscu
35	649.5	37.8	825	28	AF156720	AF156720 AF156720
36	640	37.3	942	29	AY402281	AY402281 Mus muscu
37	630	36.7	782	29	CC534047	CC534047 CH240_412
38	609.5	35.5	795	28	BZ256668	BZ256668 CH230-334
39	604.5	35.2	859	28	BZ221157	BZ221157 CH230-312
40	602.5	35.1	927	29	AY401477	AY401477 Mus muscu
41	602	35.0	758	29	CE341239	CE341239 tigr-gss-
42	601.5	35.0	669	29	CC569584	CC569584 CH240_444
43	601.5	35.0	936	29	AY402284	AY402284 Mus muscu
44	595	34.6	939	29	AY402279	AY402279 Homo sapi
45	592	34.5	786	29	AY414230	AY414230 Mus muscu

ALIGNMENTS

RESULT 1
AY401474
LOCUS
DEFINITION Mus musculus HCM0902 gene, VIRTUAL TRANSCRIPT, partial sequence,
ACCSSION AY401474
VERSION AY401474.1 GI:39757463
KEYWORDS GSS.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 948)

Qy	189	CysGluHisMetAlaValValVallysLeuAlaCysGlyAspThrArgProAsnArgValTyr	208
Db	538	TGTGAGCACATGCGCTGTGGTGAAGCTGGCTTGGGGGGGCACACAAACAACATATAT	537
Qy	209	GlyLeuThrAlaAlaLeuLeuValIleGlyValAspLeuPheCysIleGlyLeuSerTyr	228
Db	598	GGCATACAGCAGCCACACTGGTGGTAGGAAGTACTGCCATTTGTATTGCCATATCCTAT	657
Qy	229	AlaLeuIleAlaGlnAlaValLeuLeuArgLeuSerSerHisGluAlaArgSerLysAlaLeu	248
Db	658	GCACCTCATCTCCGGCGGTGTGTGGGCGCTCTCTCCAAAGAGCGCCAGGCAAGACCTTT	717
Qy	249	GlyThrCysGlySerHisValCysValIleLeuIleSerTyrThrProAlaLeuPheSer	268
Db	718	GGCACTTGTGGCTCCCACTGGGTGTGATCTCTCTCTATACACAGGGCTCTCTCA	777
Qy	269	PhePheThrHisArgPheGlyHisHisValProValHisIleHisIleLeuLeuAlaAsn	288
Db	778	TTCTATACCCAGCGCTTTGGTCAACATGTACCCAGGCAGCTCCACATCTCTGGCTGAC	833
Qy	289	ValTyrIleLeuLeuProProAlaLeuAsnProValValTyrGlyValLysThrLysGln	308
Db	838	CTCTACCTCTGTCGGCGCCATCTCTCAACCCCATCATCTATGGATGAAACCCAGCAG	897
Qy	309	IleArgLysArgValValArgValPheGlnSerGly	320
Db	898	ATCCGAGATGGAGCACTCCGACTACTCAAGAGGGGT	933
RESULT 2	AY401472	948 bp DNA linear	GSS 12-DEC-2003
LOCUS	AY401472	Homo sapiens HCM0902 gene, VIRTUAL TRANSCRIPT, partial sequence,	
DEFINITION	AY401472	genomic survey sequence.	
ACCESSION	AY401472.1	GI:39757461	
VERSION		GSS.	
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
FEATURES			
source			
gene			
ORIGIN			
Alignment Scores:			
Pred. No.:	9,66e-100	Length:	948
Score:	961.50	Matches:	178
Percent Similarity:	74.04%	Conservative:	53
Best Local Similarity:	57.05%	Mismatches:	80

```
Query Match: 55.97% Indels: 1
DB: 29 Gaps: 1
US-10-081-775-2 (1-329) x AY401478 (1-948)

QY 9 MetGluSerProHisHisThrAspValAspProSerValPhePheLeuLeuGlyIlePro 28
DB 1 ATGAAACTCATAAACCATGACCATCAGAACCAACCTCTTCTGCTCATGGGAATTC 60
QY 29 GlyLeuGluGlnPheHisLeuThrPheLeuSerLeuProValCysGlyLeuGlyThrAlaThr 48
DB 61 GCCCGGAGGCATCCACATTTGGATTGCTTTCCCTCTGCTCCATGATGCTCGGCA 120
QY 49 IleValGlyAsnIleThrIleLeuValValAlaThrGluProValLeuHisLysPro 68
DB 121 GTGCTGGGAAACATGCTGTGCTGTGCTAGTGTGCTACATTCAGAGCTGTATGTCACAGCCC 180
QY 69 ValTyrLeuPheLeuCysMetLeuSerThrIleAspLeuAlaAlaSerValSerThrVal 88
DB 181 ATGTACCTGTTCTCTGCTGCTATCCACCATGACCTGTCTCTGCACTCCACTGTG 240
QY 89 ProLysLeuAlaIlePheThrPysGlyAlaGlyHisIleSerAlaSerAlaCysLeu 108
DB 241 CCCAGCTCCTTGCACATTTTGGGCAAGGATGCTGAGATCAACTTTTGGGCGCTGTGCT 300
QY 109 AlaHisMetPheIleHisAlaPheCysMetGluSerThrValLeuAlaMet 128
DB 301 GCCCAGATGTTCTTATCCATGCTCTCAGCTGTAGAACTGTGATATCTGCTAGCAATG 360
QY 129 AlaPheAspArgTyrValAlaIleCysHisProLeuArgTyrAlaThrIleLeuThrAsp 148
DB 361 GCTTTTGACCGCTACTAGCCATTTGCTGGCTCTGCACTATGGTCTGCTCCCA 420
QY 149 ThrIleIleAlaHisIleGlyValAlaValAlaValArgGlySerLeuLeuMetLeuPro 168
DB 421 GAGTCTGTAGGCAAGCTGGGGCTGCAGCCGCTCTGCTGCTTGGGACTCATGACCCCA 480
QY 169 CysProPhePheIleGlyArgLeuAsnPheCysGlnSerHisValIleLeuHisThrTyr 188
DB 481 CTCACCTGCTTACTGGCAGACTGAGCTACTGC---AGTCGAGTGTGGCCCACTCCCTAC 537
QY 189 CysGluHisMetAlaValValLysLeuAlaCysGlyAspThrArgProAsnArgValTyr 208
DB 538 TGTGAACACATGGCTGTGCTAAAGCTGCTTGTGGAGGAACACAGCCCAACATCTAT 597
QY 209 GlyLeuThrAlaAlaLeuValIleGlyValAspLeuPheCysIleGlyLeuSerTyr 228
DB 598 GGCATCACTGCTGCCACACTGGTGTGGGCACTGACTCCATCTGTATTGCTGTCTCTAT 657
QY 229 AlaLeuIleAlaGlnAlaValLeuArgLeuSerHisGlnAlaAlaArgSerLysAlaLeu 248
DB 658 GCATCATCTCCGAGCTGTGTAGGCTTTCTCCAGGAGGCAAGGCTTAAGACCTTT 717
QY 249 GlyThrCysGlySerHisValCysValIleLeuLeuIleSerTyrThrProAlaLeuPheSer 268
DB 718 GGCATTTGTGGCTCCCACTGGGTGTATCTTCTTCTTACACACAGGACTCTTCTCC 777
QY 269 PhePheThrHisArgPheGlyHisHisValProValHisIleHisIleLeuLeuAlaAsn 288
DB 778 TTCTACACACAGCGGTGTGGCCACACAGCTGCGCCGACATCCACATCTCTTCTAGCTGAC 837
QY 289 ValTyrLeuLeuLeuProProAlaLeuAsnProValValTyrGlyValLysThrLysGln 308
DB 838 CTCTACCTGTTGTGGCCACCATCTCAACCCCATCATCTTATGTCATGAGCAACAAACAG 897
QY 309 IleArgLysArgValValArgValPheGlnSerGly 320
DB 898 ATCTGGGATGGGGCCCTTCGCGCTTCTGAAGTGGGGC 933

RESULT 3
AY401478
LOCUS Homo sapiens HCM0904 936 bp DNA linear GSS 12-DEC-2003
DEFINITION Homo sapiens HCM0904 gene, VIRUAL TRANSCRIPT, partial sequence,
```

```
genomic survey sequence.
AY401478
AY401478.1 GI:39757467
GSS.
Homo sapiens (human)
ORGANISM
REFERENCE
AUTHORS
Clark, A.G., Gnanowski, S., Nielson, R., Thomas, P., Kejariwal, A.,
Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B.,
Ferriera, S., Wang, G., Zheng, X.H., White, T.J., Shinsky, J.J.,
Adams, M.D. and Cargill, M.
TITLE
Inferring nonneutral evolution from human-chimp-mouse orthologous
gene trios
JOURNAL
Science 302 (5652), 1960-1963 (2003)
PUBMED
14671302
REFERENCE
AUTHORS
Clark, A.G., Gnanowski, S., Nielson, R., Thomas, P., Kejariwal, A.,
Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B.,
Ferriera, S., Wang, G., Zheng, X.H., White, T.J., Shinsky, J.J.,
Adams, M.D. and Cargill, M.
TITLE
Direct Submission
JOURNAL
Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
COMMENT
This sequence was made by sequencing genomic exons and ordering
them based on alignment.
FEATURES
source
1..936
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
<1..>936
/locus_tag="HCM0904"
Alignment Scores:
Pred. No.: 4.05e-93 Length: 936
Score: 904.00 Matches: 169
Percent Similarity: 71.70% Conservative: 54
Best Local Similarity: 54.34% Mismatches: 82
Query Match: 52.62% Indels: 6
DB: 29 Gaps: 2
US-10-081-775-2 (1-329) x AY401478 (1-936)
QY 7 HisAsnMetGluSerProHisHisThrAspValAspProSerValPhePheLeuLeuGly 26
DB 10 CACAAACAGACCATATTTTCAC-----CCAGTCACATTTTCTCATTCGA 54
QY 27 IleProGlyLeuGluGlnPheHisLeuThrIleLeuValValAlaThrGluProValLeuHis 46
DB 55 ATCCAGGCTCTGGAAGACTTCCACATGTGGATCTCCGGGCTTCTGCTGTGTACCTT 114
QY 47 AlaThrIleValGlyAsnIleThrIleLeuValValAlaThrGluProValLeuHis 66
DB 115 GTGCTTTGTGGCAATGCCACCATTTCTGTAGTCATCAAGGTAGAACAGACTCTCGG 174
QY 67 LysProValTyrLeuPheLeuCysMetLeuSerThrIleAspLeuAlaAlaSerValSer 86
DB 175 GAGCCCATGTTCTTCTCCGGCCATCTTTCCACTATTGATTTGGGCTTCTTCTGCAACC 234
QY 87 ThrValProLysLeuLeuAlaIlePheTyrCysGlyAlaGlyHisIleSerAlaSerAla 106
DB 235 TCTGTGCTCGCATGTGGGTATCTTCTGTTGATGCTCAGGAGATATCATGTGAGCT 294
QY 107 CysLeuAlaHisMetPhePheIleHisAlaPheCysMetMetGluSerThrValLeuLeu 126
DB 295 TGTGTGCCCAGATGTTTCTGATCCATGCTTCACTGGCATGGAGGCTGAGGCTTACTG 354
QY 127 AlaMetAlaPheAspArgTyrValAlaIleCysHisProLeuArgTyrAlaThrIleLeu 146
DB 355 GCTATGCTTTTGACCGTTAATGTGGCCATCTGTGCTCCACTACATTACGCAACCATCTTG 414
```

```

147 ThrAspThrIleAlaHisIleGlyValAlaValArgGlySerLeuLeuMet 166
148 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
149 ACATCCCTAGTGTGGGCGCATAGCATGTCATTTGTAATTCGTCCTGTTTACTTACA 474
150 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
151 LeuProCysProPheIleGlyArgLeuAsnPheCysGlnSerHisValIleLeuHis 186
152 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
153 CTTCCTCCATGGTCTATCTTATCTACCGCTACCTTTTGTCCAGGCTCACAATAGCCCAT 534
154 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
155 ThrTyrCysGluHisMetAlaValIleValIleValIleValIleValIleValIle 206
156 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
157 TCCTACTGTGAGCAGCATGGGCGATTGCAAAATGCTCTGTGAAACATTCGTATCAATGGT 594
158 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
159 ValTyrGlyLeuThrAlaAlaLeuLeuValIleGlyValAlaLeuPheCysIleGlyLeu 226
160 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
161 ATCTATGGGCTTTTGTAGTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTT 651
162 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
163 SerTyrAlaLeuIleAlaGlnAlaValLeuArgLeuSerHisGluAlaArgSerLys 246
164 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
165 TCGTATGTTTACATCTCCGGTGTCTTCCGCTTCCCATCAGATGCTCAGCTAATA 711
166 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
167 AlalaLeuGlyThrCysGlySerHisValCysValIleLeuIleSerTyrThrProAlaLeu 266
168 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
169 GCCTAACACGCTGGGCGCTCATGTTGGAGTCATCTGTCTTTCTATATCCCTCAGTC 771
170 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
171 PheSerPheThrHisArgPheGlyHisValProValHisIleHisIleLeuLeu 286
172 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
173 TTCTCTTCTTCTTACTCATCGATTTGGACACAAATACCAGGTTACATTCACATCTTCTTT 831
174 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
175 AlalaLeuValTyrLeuLeuProAlaLeuAsnProValValTyrGlyValLysThr 306
176 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
177 GCCAATCTCTATTTGATTTATCCACCTCTCTCAACCCCATCATTTATGGGTGAGGACC 891
178 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
179 LysGlnIleArgLysArgValValArgValPhe 317
180 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
181 AAACAGATTCGAGCGAGTGCTCATGTTT 924
182 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 4
AY401473 948 bp DNA linear GSS 12-DEC-2003
LOCUS Pan troglodytes HCM0902 gene, VIRTUAL TRANSCRIPT, partial sequence,
DEFINITION genomic survey sequence.
ACCESSION AY401473
VERSION AY401473.1 GI:39757462
KEYWORDS GSS.
SOURCE Pan troglodytes (chimpanzee)
ORGANISM Pan troglodytes
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
1 (bases 1 to 948)
Clark,A.G., Gnanowski,S., Nielson,R., Thomas,P., Kejarawal,A.,
Todd,M.A., Tanenbaum,D.M., Civallo,D.R., Lu,F., Murphy,B.,
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
Inferring nonneutral evolution from human-chimp-mouse orthologous
gene trios
Science 302 (5652), 1960-1963 (2003)
14671302
2 (bases 1 to 948)
Clark,A.G., Gnanowski,S., Nielson,R., Thomas,P., Kejarawal,A.,
Todd,M.A., Tanenbaum,D.M., Civallo,D.R., Lu,F., Murphy,B.,
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
Direct Submission
Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
This sequence was made by sequencing genomic exons and ordering
them based on alignment
location/Qualifiers
1..948
/organism="Pan troglodytes"
/mol_type="genomic DNA"
/db_xref="taxon:9598"
FEATURES
source

```

```

gene <1..>948
/locus_tag="HCM0902"
ORIGIN
Alignment Scores:
Pred. No.: 1.94e-91 Length: 948
Score: 889.50 Matches: 169
Percent Similarity: 70.19% Conservative: 50
Best Local Similarity: 54.17% Mismatch: 92
Query Match: 51.78% Indels: 1
DB: 29 Gaps: 1
US-10-081-775-2 (1-329) x AY401473 (1-948)
QY 9 MetGluSerProHisHisThrAspValAspProSerValPheLeuLeuGlyIlePro 28
DB 1 ATGAACCTCATAAACCATAGCCATCAGAACCAACCTCTTCTGCTCATGGGAATCCA 60
QY 29 GlyLeuGluGlnPheHisLeuTyrLeuSerLeuProValCysGlyLeuGlyThrAlaThr 48
DB 61 GGCCTGGAGGCATCCCACTTTTGGATTGCTTTCCCTTCTGCTCATGTATGCCCTGGCA 120
QY 49 IleValGlyAsnIleThrIleLeuValValAlaThrGluProValIleHisLysPro 68
DB 121 GTGCTGGGAACATGTTGGTGTCTGCTAGTGTATTCAGAGCCTGTATTCAGCAGCCC 180
QY 69 ValTyrLeuPheLeuCysMetLeuSerThrIleAspLeuAlaAlaSerValSerThrVal 88
DB 181 ATGATACCTGTTCTTCTGCTATCCACCATTCACCTGGTCTCTGCACCTCCACTGTG 240
QY 89 ProLysLeuLeuAlaIlePheTyrCysGlyAlaGlyHisIleSerAlaSerAlaCysLeu 108
DB 241 CCCAAGCTCTTGCATTTTGGGCAAGGATGCTGAGATCAACTTTGGGCGCTGTGCT 300
QY 109 AlaHisMetPhePheIleHisAlaPheCysMetMetGluSerThrValLeuLeuAlaMet 128
DB 301 GCCAGATGTTCTTTATCCATGGCTTCTCAGCTAGAAATCTGTTATCTGCTAGCAATG 360
QY 129 AlaPheAspArgTyrValAlaIleCysHisProLeuArgTyrAlaThrIleLeuThrAsp 148
DB 361 GCCTTTGACCGCTACTTATAGCCATCTCGCGGCTCTGCACTATGGGTCAATGCTCTCCA 420
QY 149 ThrIleIleAlaHisIleGlyValAlaValAlaValArgGlySerLeuLeuMetLeuPro 168
DB 421 GAGTCTGTGNGCAGCTGGGGCTGCAGCCNTGCTCGNGGTTTGGGACTCATGACCCCA 480
QY 169 CysProPhePheIleGlyArgLeuAsnPheCysGlnSerHisValIleHisThrTyr 188
DB 481 CTCACCTGCTTACTGGCAAGACTGNNNTAC---TNNNGTCTNNNNNGGCCCTCTAC 537
QY 189 CysGluHisMetAlaValValLysLeuAlaCysGlyAspThrArgProAsnArgValTyr 208
DB 538 TNNNAACATGCTGCTGTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 597
QY 209 GlyLeuThrAlaAlaLeuLeuValIleGlyValAlaLeuPheCysIleGlyLeuSerTyr 228
DB 598 GGCATCAGCGCTGCCACACTGNNNNNGGCACTGACTCCATCTGTTATGCTGCTCTAT 657
QY 229 AlaLeuIleAlaGlnAlaValLeuArgLeuSerHisGluAlaArgSerLysAlaLeu 248
DB 658 GCACATCATCTCCGGGCTGTGTAGGCTTTCTTCCAGAGGCGAGGCTTAGACCTTT 717
QY 249 GlyThrCysGlySerHisValCysValIleLeuSerTyrThrProAlaLeuPheSer 268
DB 718 GGCATTGTGGTCCACCTGGGTGTCATCTCTCTTCTACACACAGGACTCTTCTCC 777
QY 269 PhePheThrHisArgPheGlyHisValProValHisIleHisIleLeuLeuAlaAsn 288
DB 778 TTCTACACACAGCGGTTTGGCAGCAGCTGCCCGCTCAGATCCATCTCTTAGCTGAC 837
QY 289 ValTyrLeuLeuLeuProAlaLeuAsnProValValTyrGlyValLysThrLysGln 308
DB 838 CTCCTACCTGGTGTGCCACCCCATGCTCAACCCCATCATCTATGCGCATGAGACCAACAG 897

```

```

QY 309 IleArgLysArgValValValArgValPheGlnSerGly 320
|||
Db 898 ATCTGGATGGGGCCCTCCGGCTTCGAAGTGGGGC 933
|||
RESULT 5
AY414225
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS
Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejariwal, A.,
Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B.,
Ferreira, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J.,
Adams, M.D. and Cargill, M.
TITLE
Inferring nonneutral evolution from human-chimp-mouse orthologous
gene trios
JOURNAL
Science 302 (5652), 1960-1963 (2003)
PUBMED
14671302
REFERENCE
AUTHORS
Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejariwal, A.,
Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B.,
Ferreira, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J.,
Adams, M.D. and Cargill, M.
TITLE
Direct Submission
JOURNAL
Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
COMMENT
This sequence was made by sequencing genomic exons and ordering
them based on alignment.
FEATURES
source
location/Qualifiers
1..957
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
<1..>957
/locus_tag="HCM5153"
gene
ORIGIN
Alignment Scores:
Pred. No.: 1,97e-91 Length: 957
Score: 889.50 Matches: 160
Percent Similarity: 69.41% Conservative: 51
Best Local Similarity: 52.63% Mismatches: 92
Query Match: 51.78% Indels: 1
DB: 29 Gaps: 1
US-10-081-775-2 (1-329) x AY414225 (1-957)
QY 19 ProSerValPhePheLeuLeuGlyLeuProGlyLeuGluGlnPheHisLeuTyrLeuSer 38
|||
Db 34 CCAGACACCTTCTTCTTAACAGGATCCAGGGTGGAGGCTGCCACCTTCGGATTGCC 93
|||
QY 39 LeuProValCysGlyLeuGlyThrAlaThrIleValGlyAsnIleThrIleLeuValVal 58
|||
Db 94 ATCCCTTCTGTCGCGCATGATCTTGTAGCACTGTTGGAAATGCTGCCCTCATCTGGTC 153
|||
QY 59 ValAlaThrGluProValLeuHisLysProValTyrLeuPheLeuCysMetLeuSerThr 78
|||
Db 154 ATTCCATGGCAATAGCTCTTCATGACCTATGACCTCTTCCTGCTGCTTCTCTCACTC 213
|||
QY 79 IleAspLeuAlaAlaSerValSerThrValProLysLeuLeuAlaIlePheTyrCysGly 98
|||
Db 214 ACAGACTGCTCTCAGTTCTACCACTGTGCCAAGATGCTGGCCATTTGTGGCTCCAT 273
|||
QY 99 AlaGlyHisIleSerAlaSerAlaCysLeuAlaHisMetPhePheIleHisAlaPheCys 118
|||

```

```

Db 274 GCTGGTGAGATTCTCTTGGTGGATGCGTGGCCGAGATGTTTTGTGTCATCTTATCTAT 333
QY 119 MetMetGluSerThrValLeuLeuAlaMetAlaPheAspArgTyrValAlaIleCysHis 138
|||
Db 334 GCTCTGGAGTCTCGATTCTACTTGCATGGCTTTGATAGGTATGTGGTATCTGTAAAC 393
|||
QY 139 ProLeuArgTyrAlaThrIleLeuThrAspThrIleIleAlaHisIleGlyValAlaAla 158
|||
Db 394 CATTAAAGTATACAAACCATTTCTCAACCATGCTGTATAGGAGGAGATGGCTTTGTGG 453
|||
QY 159 ValValArgGlySerLeuLeuMetLeuProCysProPhePheIleGlyArgLeuAsnPhe 178
|||
Db 454 CTATTCGCTAGTGGCTATGCTCTCCCTTCATCTTCTGCTGAGGAGATCCCTCTAC 513
|||
QY 179 CysGlnSerHisValIleLeuHisThrTyrCysGluHisMetAlaValValValValLeuAla 198
|||
Db 514 TGTGGTCAACCGTGTGATGACACACATACCTGTGAGCATATGGGCATCGCCGATGGCC 573
|||
QY 199 CysGlyAspThrArgProAsnArgValTyrGlyLeuThrAlaAlaLeuLeuValIleGly 218
|||
Db 574 TGTGCCAACATCACTGTCTCAATATTGCTATGGGCTTAACCTGTGGCTCTGTGGCCATGG 633
|||
QY 219 ValAspLeuPheCysIleGlyLeuSerTyrAlaLeuIleAlaGlnAlaValLeuArgLeu 238
|||
Db 634 CTGGATTCATCTCATTTGCCATTTCTTATGGCTTTATCTCCATGCGAGTCTTTTACCTT 693
|||
QY 239 SerSerHisGluAlaArgSerLysAlaLeuGlyThrCysGlySerHisValCysValIle 258
|||
Db 694 CCATCTCATGATGCCAGCACACAAAGCTCTGAGTACCTGTGGCTCCACCATTTGGCATCATC 753
|||
QY 259 LeuIleSerTyrThrProAlaLeuPheSerPheThrHisArgPheGlyHisHis--- 277
|||
Db 754 CTGGTTTTCTACATCCCTGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 813
|||
QY 278 ValProValHisIleHisIleLeuAlaAsnValTyrLeuLeuLeuProProAlaLeu 297
|||
Db 814 GTCCCCAAGCATGTGCATCTTCTGCTTAATCTCTAATGCTGCTGCTGCTGCTGCTGCTGCT 873
|||
QY 298 AsnProValValTyrGlyValLysThrLysGlnIleArgLysArgValValArgValPhe 317
|||
Db 874 AATCCTATTCTATGGAGCTAGAACCAAGGAGATTCGGAGTCGAGCTTCTAAACATGCTT 933
|||
QY 318 GlnSerGlyGln 321
|||
Db 934 CACCTGGGGAAG 945
|||
RESULT 6
AY414227
LOCUS
DEFINITION
Mus musculus HCM5153 gene, VIRTUAL TRANSCRIPT, partial sequence,
genomic survey sequence.
ACCESSION
AY414227
VERSION
AY414227.1 GI:39770189
KEYWORDS
GSS.
SOURCE
Mus musculus (house mouse)
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
AUTHORS
Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejariwal, A.,
Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B.,
Ferreira, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J.,
Adams, M.D. and Cargill, M.
TITLE
Inferring nonneutral evolution from human-chimp-mouse orthologous
gene trios
JOURNAL
Science 302 (5652), 1960-1963 (2003)
PUBMED
14671302
REFERENCE
AUTHORS
Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejariwal, A.,
Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B.,
Ferreira, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J.,
Adams, M.D. and Cargill, M.
TITLE
Direct Submission

```

Db 751 CTGTTTCTATATACCTGCTTCTTCTCTCTCTACCCCATCGCTTGGTCACAATCGA 811
 Qy 278 ValProValHisIleHisIleLeuLeuAlaAsnValTyrLeuLeuLeuProAlaIleu 297
 Db 811 GTTCCCAAGCATGTGCATATCTTCTGCTAACTCTATGCTGGTGGCTCCCGTGCTC 870
 Qy 298 AsnProValValTyrGlyValLysThrLysGlnIleArgLysArgValValArgValPhe 317
 Db 871 AATCCCATCATCTACGCTGCAAGAACCAAGAGATTCGGAGTCGGCTTCTAAAACTGCTT 930
 Qy 318 GlnSerGlyGlnGlyMet 323
 Db 931 CACTTGGGAAGGACTTGA 948

 RESULT 7
 AY401480
 LOCUS Mus musculus HCM0904 gene, VIRTUAL TRANSCRIPT, partial sequence,
 DEFINITION genomic survey sequence.
 ACCESSION AY401480
 VERSION AY401480.1 GI:39757469
 KEYWORDS GSS.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1 (bases 1 to 936)
 AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarawal,A.,
 Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
 Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
 Adams,M.D. and Cargill,M.
 TITL Inferring nonneutral evolution from human-chimp-mouse orthologous
 gene trios
 JOURNAL Science 302 (5652), 1960-1963 (2003)
 PUBMED 146711302
 REFERENCE 2 (bases 1 to 936)
 AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarawal,A.,
 Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
 Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
 Adams,M.D. and Cargill,M.
 TITL Direct Submission
 JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
 Rockville, MD 20850, USA
 COMMENT This sequence was made by sequencing genomic exons and ordering
 them based on alignment.
 FEATURES
 source Location/Qualifiers
 1..936
 /organism="Mus musculus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10090"
 <1..>936
 /locus_tag="HCM0904"
 gene
 ORIGIN
 Alignment Scores:
 Pred. No.: 2,086-90 Length: 936
 Score: 880.50 Matches: 163
 Percent Similarity: 70.42% Conservative: 56
 Best Local Similarity: 52.41% Mismatches: 87
 Query Match: 51.25% Indels: 5
 DB: 29 Gaps: 2

 US-10-081-775-2 (1-329) x AY401480 (1-936)
 Qy 7 HisAsnMetGluSerProHisHisThrAspValAspProSerValPhePheLeuLeuGly 26
 Db 7 CATAATAATAGAGCGTTCTCTCAC-----CCAGATACATATTTTCTCTCATGGG 54
 Qy 27 IleProGlyLeuGluGlnPheHisIleLeuTrpLeuSerLeuProValCysGlyLeuGlyThr 46
 Db 55 ATCCCAAGGCTCGAGAAATCATCGCTGGATTTCTCTGCCCTTCTGCTGATTATCTT 114
 Qy 47 AlaThrIleValGlyAsnIleThrIleLeuValValAlaThrGluProValLeuHis 66

Db 115 GTGGCCCTAATGTAATACCATGATCTGGTAGTATTATCAAGACAGCGTCTCTCCGA 174
 Qy 67 LysProValTyrLeuPheLeuCysMetLeuSerThrIleAspLeuAlaSerValSer 86
 Db 175 GAGCCCATGTTCTTACTTCTGGCCATCTCTTCTCAGTCGACTGGCTCTTCTTACAACC 234
 Qy 87 ThrValProLysLeuLeuAlaIlePheTrpCysGlyAlaGlyHisIleSerAlaSerAla 106
 Db 235 TCTGTGCCCGTATGCTTGGCATATTCTGTGTGATGCCCATGAATAAATCTCGGTGGC 294
 Qy 107 CysLeuAlaHisMetPhePheIleHisAlaPheCysMetMetGluSerThrValLeuLeu 126
 Db 295 TGTGTGGCTCAGATGTTTTGTATCCATGCTTTACTGGCATGGAGCTGAGGTCTCATG 354
 Qy 127 AlaMetAlaPheAspArgTyrValAlaIleCysHisProLeuArgTyrAlaThrIleLeu 146
 Db 355 GCCATGGCTTTTGACCGTTATGTAGCCATCTGTGTCTCCACTCCACTACACACCATCTTG 414
 Qy 147 ThrAspThrIleAlaHisIleGlyValAlaAlaValAlaValArgGlySerLeuLeuMet 166
 Db 415 ACATCTGTGCTCATGGCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 474
 Qy 167 LeuProCysProPhePheIleGlyArgLeuAsnPheCysGlnSerHisValIleLeuHis 186
 Db 475 TGCCCAATGGTCTATCTCATTTACCGTTTACCGTTTGTCAAGCTCATGATATGCCCAT 534
 Qy 187 ThrTyrCysGluHisMetAlaValValIleValAlaCysGlyAspThrArgProAsnArg 206
 Db 535 TCCTATTGTGAACATGCGCATGCGCAATGTCCTGTGGAGACATCCATCAACGCT 594
 Qy 207 ValTyrGlyLeuThrAlaAlaLeuValIleGlyValAspLeuPheCysIleGlyLeu 226
 Db 595 GTGATGGCTTTTGTGTCT 651
 Qy 227 SerTyrAlaLeuAlaGlnAlaValLeuArgLeuSerSerHisGluAlaArgSerLys 246
 Db 652 TCATATGGTATCATCTCTCGTCTGTGTGCTCCCTCCATCTCAGATGACGGTTTAA 711
 Qy 247 AlaLeuGlyThrCysGlySerHisValCysValIleLeuIleSerTyrThrProAlaLeu 266
 Db 712 GCCCTCAGACATGTTGGTTCCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 771
 Qy 267 PheSerPheThrHisArgPheGlyHisHisValProValHisIleHisIleLeuLeu 286
 Db 772 TTCCTTCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCT 831
 Qy 287 AlaAsnValTyrLeuLeuProProAlaLeuAsnProValValTyrGlyValLysThr 306
 Db 832 GCCAACCTCTATTGGTCTATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 891
 Qy 307 LysGlnIleArgLysArgValValArgValPhe 317
 Db 892 AAGCAGATAAGAGAATGTTCTCCACATATTC 924

AK036356 1964 bp mRNA linear HTC 19-SEP-2003
 Mus musculus 16 days neonate cerebellum cDNA, RIKEN full-length
 enriched library, clone:9630060D12 product:OLFACTORY RECEPTOR
 MOR18-2, full insert sequence.
 AK036356
 AK036356-1 GI:26331331
 HTC; CAP trapper.
 Mus musculus (house mouse)
 Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 1
 Carninci, P. and Hayashizaki, Y.
 High-efficiency full-length cDNA cloning
 JOURNAL Meth. Enzymol. 303, 19-44 (1999)
 MEDLINE 99279253

10349636
 2 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Sasaki, N., Carninci, P.,
 Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
 Normalization and subtraction of cap-trapper-selected cDNAs to
 prepare full-length cDNA libraries for rapid discovery of new genes
 Genome Res. 10 (10), 1617-1630 (2000)
 20499374
 11042159
 3 Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
 Konno, H., Akiyama, J., Nishi, K., Katsunai, T., Tashiro, H., Itoh, M.,
 Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,
 Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
 Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M.,
 Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J.,
 Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
 RIKEN integrated sequence analysis (RISA) system--384-format
 sequencing pipeline with 384 multicapillary sequencer
 Genome Res. 10 (11), 1757-1771 (2000)
 20530913
 11076861
 4 The RIKEN Genome Exploration Research Group Phase II Team and the
 FANTOM Consortium.
 Functional annotation of a full-length mouse cDNA collection
 Nature 409, 685-690 (2001)
 5 The FANTOM Consortium and the RIKEN Genome Exploration Research
 Group Phase I & II Team.
 Analysis of the mouse transcriptome based on functional annotation
 of 60,770 full-length cDNAs
 Nature 420, 563-573 (2002)
 6 (bases 1 to 1964)
 Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P.,
 Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W.,
 Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T.,
 Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T.,
 Kato, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,
 Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,
 Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N.,
 Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N.,
 Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T.,
 Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S.,
 Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A.,
 Muramatsu, M. and Hayashizaki, Y.
 Direct Submission
 Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of
 Physical and Chemical Research (RIKEN), Laboratory for Genome
 Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
 RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
 Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.go.jp,
 URL: http://genome-gsc.riken.go.jp/, Tel: 81-45-503-9222,
 Fax: 81-45-503-9216)
 cDNA library was prepared and sequenced in Mouse Genome
 Encyclopedia Project of Genome Exploration Research Group in Riken
 Genomic Sciences Center and Genome Science Laboratory in RIKEN.
 Division of Experimental Animal Research in Riken contributed to
 prepare mouse tissues.
 Please visit our web site for further details.
 URL: http://genome.gsc.riken.go.jp/
 URL: http://fantom.gsc.riken.go.jp/
 Location/Qualifiers
 1. 1964
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6J"
 /db_xref="FANTOM_DB:9630060D12"
 /db_xref="MGI:2400778"
 /db_xref="taxon:10090"
 /clone="9630060D12"
 /tissue_type="cerebellum"
 /clone_lib="RIKEN full-length enriched mouse cDNA library"


```

DEFINITION Mus musculus ORS2R1 gene, VIRTUAL TRANSCRIPT, partial sequence,
genomic survey sequence.
ACCESSION AY410598
VERSION AY410598.1 GI:39766568
KEYWORDS GSS.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous
gene trios
JOURNAL Science 302 (5652), 1960-1963 (2003)
PUBMED 14671302
REFERENCE
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE Direct Submission
JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
COMMENT This sequence was made by sequencing genomic exons and ordering
them based on alignment.
FEATURES
source
1..948
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
<!.>.948
/gene="ORS2R1"
/locus_tag="HCM3944"
ORIGIN
Alignment Scores:
Pred. No.: 6,696-89 Length: 948
Score: 867.50 Matches: 168
Percent: 67.0% Conservative: 42
Best Local Similarity: 53.6% Mismatches: 98
Query Match: 50.4% Indels: 5
DB: 29 Gaps: 1
US-10-081-775-2 (1-329) x AY410600 (1-948)
QY 5 LeuGlyHisAsnMetGluSerProHisHisThrAspValAspProSerValPheLeu 24
Db 7 TTGGCTTCAGGAACAGCTCTCCCACTCTACA-----TTTTTCATCTTG 51
QY 25 LeuGlyIleProGlyLeuGlnPheHisLeuTrpLeuSerLeuProValCysGlyLeu 44
Db 52 CTTGGAATCCAGGACTGGAGAATTACCAATTCGGTTGCTTCCATCTGTGTGTCATG 111
QY 45 GlyThrAlaThrIleValGlyAsnIleThrIleLeuValValAlaThrGluProVal 64
Db 112 TATATTGTGGCAGTGACTGGAAACATCACCATCTCTTCACATATCCGAATGACCACAC 171
QY 65 LeuHisLysProValTrpLeuPheLeuCysMetLeuSerThrIleAspLeuAlaAsp 84
Db 172 CTAATGAGCCCATGATCTCTTCTGGCGATGCTGGCTACCACTGACCTTGGCTGCTCC 231
QY 85 ValSerThrValProLysLeuLeuAlaIlePheTrpCysGlyAlaGlyHisLeuSerAla 104
Db 232 TCCTCCACACACACCAAAATGTGGCTATACCTGCTTTCATGATCATGAAGATTGAATAT 291
QY 105 SerAlaCysLeuAlaHisMetPhePheIleHisAlaPheCysMetMetSerThrVal 124
Db 292 CATGCTGCTCTATCAGGTGTCTTCATACATGCTTCTTCTGCTGGAGTCTGGGTG 351
QY 125 LeuLeuAlaMetAlaPheAspArgTrpValAlaIleCysHisProLeuArgTrpAlaThr 144

```

```

Db 352 CTCATGACCATGGCGCTTGGACCGTTATGTGCTATCTGCTTCCCACTCCGACATTCACGC 411
QY 145 IleLeuThrAspThrIleAlaHisIleGlyValAlaAlaValValArgGlySerLeu 164
Db 412 ATCTCCACCATCTGCAGTCATCAAACTAGTGTCTGTGTGTGAGGGGTGCTG 471
QY 165 LeuMetLeuProCysProPhePheIleGlyArgLeuAsnPheCysGlnSerHisValIle 184
Db 472 TGGGTGAGCCCTTCTGTTCATGCTCTCCAGGATGCCCTTCTGCCCAACAAGATCAT 531
QY 185 LeuHisThrTrpCysGluHisMetAlaValValLysLeuAlaCysGlyAspThrArgPro 204
Db 532 CCCAGTCTCTACTGTGAGCACATGGCTGTCTTAAGTTGGTATGTCTGTATACCAAGATC 591
QY 205 AsnArgValTrpGlyLeuThrAlaLeuLeuValIleGlyValAlaAspLeuPheCysIle 224
Db 592 AATCGTGATATGGACTCTTTGTGCTTCTCTCTTGTGGCTTTTGACATTAATTCATC 651
QY 225 GlyLeuSerTrpAlaLeuIleAlaGlnAlaValLeuArgLeuSerSerHisGluAlaArg 244
Db 652 AGTGATCTTATGATGATGATCTGAGAGCTGTGCTGAGTTGCCCTCAGGTGAAGCTGCG 711
QY 245 SerLysAlaLeuGlyThrCysGlySerHisValCysValIleLeuIleSerTrpThrPro 264
Db 712 CTCAGGCTTTTGGTACATGTCTCTCATATTGGTGTCTATTAACCTTATATATCCCA 771
QY 265 AlaLeuPheSerPhePheThrHisArgPheGlyHisHisValProValHisIleHisIle 284
Db 772 GCCCTTTTCACCTTCCCTCACCACCGCTTGGTACCATGTGGCCCGAGTTGTCACATC 831
QY 285 LeuLeuAlaAsnValTrpLeuLeuLeuProAlaLeuAsnProValValTrpGlyVal 304
Db 832 ATGTTTGCATGCTCTCTCTAGTCTCTCTATGCTCAATCCCATCATATATGAGTT 891
QY 305 LysThrLysGlnIleArgLysArgValValArgValPhe 317
Db 892 AGAACCAACAGATCAGACAGAGGTTTACCAGGGGTTT 930
RESULT 11
AY410598 948 bp DNA linear GSS 16-DEC-2003
LOCUS Homo sapiens ORS2R1 gene, VIRTUAL TRANSCRIPT, partial sequence,
DEFINITION Genomic survey sequence.
ACCESSION AY410598
VERSION AY410598.1 GI:39766566
KEYWORDS GSS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous
gene trios
JOURNAL Science 302 (5652), 1960-1963 (2003)
PUBMED 14671302
REFERENCE
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE Direct Submission
JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
COMMENT This sequence was made by sequencing genomic exons and ordering
them based on alignment.
FEATURES
source
1..948
/organism="Homo sapiens"

```

[illegible]

